

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model.  
Run on: May 25, 2002, 11:55:09 ; Search time 281.82 Seconds  
(without alignments)  
9083.526 Million cell updates/sec

Title: US-09-667-380a-1

Perfect score:

Sequence: I atggatcgctgtccgggggg...ggatcttgcgtccaggag 1491

Scoring table: IDENTITY\_NUC Gapext 1.0

Searched: 1736436 seqs, 88457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_032802.\*

1: /SIBS1/geodata/geneseq/geneseq/geneseq -emb1/NA1980.DAT:\*

2: /SIBS1/geodata/geneseq/geneseq -emb1/NA1981.DAT:\*

3: /SIBS1/geodata/geneseq/geneseq -emb1/NA1982.DAT:\*

4: /SIBS1/geodata/geneseq/geneseq -emb1/NA1983.DAT:\*

5: /SIBS1/geodata/geneseq/geneseq -emb1/NA1984.DAT:\*

6: /SIBS1/geodata/geneseq/geneseq -emb1/NA1985.DAT:\*

7: /SIBS1/geodata/geneseq/geneseq -emb1/NA1986.DAT:\*

8: /SIBS1/geodata/geneseq/geneseq -emb1/NA1987.DAT:\*

9: /SIBS1/geodata/geneseq/geneseq -emb1/NA1988.DAT:\*

10: /SIBS1/geodata/geneseq/geneseq -emb1/NA1989.DAT:\*

11: /SIBS1/geodata/geneseq/geneseq -emb1/NA1990.DAT:\*

12: /SIBS1/geodata/geneseq/geneseq -emb1/NA1991.DAT:\*

13: /SIBS1/geodata/geneseq/geneseq -emb1/NA1992.DAT:\*

14: /SIBS1/geodata/geneseq/geneseq -emb1/NA1993.DAT:\*

15: /SIBS1/geodata/geneseq/geneseq -emb1/NA1994.DAT:\*

16: /SIBS1/geodata/geneseq/geneseq -emb1/NA1995.DAT:\*

17: /SIBS1/geodata/geneseq/geneseq -emb1/NA1996.DAT:\*

18: /SIBS1/geodata/geneseq/geneseq -emb1/NA1997.DAT:\*

19: /SIBS1/geodata/geneseq/geneseq -emb1/NA1998.DAT:\*

20: /SIBS1/geodata/geneseq/geneseq -emb1/NA1999.DAT:\*

21: /SIBS1/geodata/geneseq/geneseq -emb1/NA2000.DAT:\*

22: /SIBS1/geodata/geneseq/geneseq -emb1/NA2001A.DAT:\*

23: /SIBS1/geodata/geneseq/geneseq -emb1/NA2001B.DAT:\*

24: /SIBS1/geodata/geneseq/geneseq -emb1/NA2002.DAT:\*

#### ALIGNMENTS

#### RESULT

1

ID AAFT7686 standard; cDNA; 1491 BP.

XX AAFT7686;

AC AAFT7686;

XX 29-MAY-2001 (first entry)

XX DE Human protease-inhibitor like protein coding sequence #1.

XX KW Human; protease-inhibitor like protein; NHP; novel human protein;

XX KW trypsin inhibitor; gene therapy; drug screening; disease diagnosis;

XX KW signal transduction; ss.

OS Homo sapiens.

XX WO200121051-A2.

PN PD 29-MAR-2001.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1491	100.0	1491 22 AAFT7686	Human protease-inh
2	1491	100.0	22 AAFT7686	Human protease-inh
3	1487.8	99.8 ✓	4877 22 [AAFT7686]	Human cancer agent✓
4	1487.8	99.8 ✓	4877 22 AAS60893	Human cancer agent✓
5	1486.2	99.7 ✓	2400 22 AAD1764	Human novel trypsin✓
6	1474.8	98.9 ✓	2403 22 AAD1764	Human novel trypsin✓
7	1452.6	97.4 ✓	2412 22 AAD1764	Human novel trypsi
8	1393.8	93.5	1669 22 AAD1765	Human novel trypsi

PT Novel isolated human protease inhibitor-like polynucleotide useful in  
 PT therapeutic, diagnostic and pharmacogenetic applications -  
 XX  
 PS  
 XN  
 C1aim 1; Page 26; 29pp; English.

CC The present invention provides the protein and coding sequences of the  
 CC mammalian trypsin inhibitors. The sequences are useful in disease  
 diagnosis and treatment, particularly of diseases associated with signal  
 transduction. The present sequence is one version of the NHP cDNA.  
 XX  
 Sequence 1491 BP; 368 A; 409 C; 426 G; 288 T; 0 other;

Query Match 100 %; score 1491; DB 22; Length 1491;  
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;  
 Matches 1491; Conservative 0; -

QY	1	at gagatgcgtctggatgtatccatccatgggtctgttctgttcggatcc 60
Db	1	at gagatgcgtctggatgtatccatccatgggtctgttctgttcggatcc 60
Qy	61	caag getac ctccaa caa ctgg tca ctcttt aagg atc gtcc aat cc aac 120
Db	61	caag getac ctccaa caa ctgg tca ctcttt aagg atc gtcc aat cc aac 120
Qy	121	aac gatgtcttactccgggtccgaggccatcccaaggaga aaggatcc 180
Db	121	aac gatgtcttactccgggtccgaggccatcccaaggaga aaggatcc 180
Qy	181	at gctq cacaca caa agtcttggggcaggatggcaggatcgatggata 240
Db	181	at gctq cacaca caa agtcttggggcaggatggcaggatcgatggata 240
Qy	241	act cggat gac gact ctggat gaga gatctgtcgatggatggatcatg 240
Db	241	act cggat gac gact ctggat gaga gatctgtcgatggatggatcatg 240
Qy	301	ca cggccacc aactctgtgggtccatggccaggatggatggatggatgg 360
Db	301	ca cggccacc aactctgtgggtccatggccaggatggatggatggatgg 360
Qy	361	tatcgctccggggttccatgtgcaggatgtcgatggatggatggatgg 360
Db	361	tatcgctccggggttccatgtgcaggatgtcgatggatggatggatgg 360
Qy	421	ccttacccggcgaatgtgcacccctgggtccaggagggttccatgtgc 480
Db	421	ccttacccggcgaatgtgcacccctgggtccaggagggttccatgtgc 480
Qy	481	cactacacac aactgttggccaccaccaaa agatcggtgtgtgaaactcg 540
Db	481	cactacacac aactgttggccaccaccaaa agatcggtgtgtgaaactcg 540
Qy	541	cggaa gactcttggaa aagtttggaa gacggcggttactttgtcgaaatt 600
Db	541	cggaa gactcttggaa aagtttggaa gacggcggttactttgtcgaaatt 600
Qy	601	tctcca aagg gaaactggatggatggatggatggatggatggatggatgg 660
Db	601	tctcca aagg gaaactggatggatggatggatggatggatggatggatgg 660
Qy	661	tccocacccgatctatggggggaaacttggatggatggatggatggatgg 720
Db	661	tccocacccgatctatggggggaaacttggatggatggatggatggatgg 720
Qy	721	actccaa aacccggaa cggaa cggaa aacttggatggatggatggatgg 780
Db	721	actccaa aacccggaa cggaa cggaa aacttggatggatggatggatgg 780
Qy	781	aacca gtttgcgtccaa accggatggatggatggatggatggatggatgg 840
Db	781	aacca gtttgcgtccaa accggatggatggatggatggatggatggatgg 840

RESULT 2  
 ID AAF77687  
 ID AAF77687 standard; cDNA; 2272 BP.  
 AC AAF77687;  
 XX  
 DR 29-MAY-2001 (first entry)  
 XX  
 DE Human protease-inhibitor like protein coding sequence #2.  
 XX  
 KW Human; protease-inhibitor like protein; NHP; novel human protein;  
 KW trypsin inhibitor; gene therapy; drug screening; disease diagnosis;  
 KW signal transduction; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200121651-A2.  
 XX  
 PD 29-MAR-2001.  
 XX  
 PP 22-SRP-2000; 2000WO-US26048.  
 XX  
 PR 24-SEP-1999; 99US-0156101.  
 XX  
 PA (LEXI-) LEXICON GENETICS INC.  
 XX  
 PI Donoho G., Turner CA, Wattler F., Nehls M., Friedrich G;











Database: N\_Genescoy\_032802

Query Match	98.9%	Score 1474.8;	DB 22;	Length 2403;	
Best Local Similarity	99.7%	Pred. No. 0:			
Matches 1489;	Conservative	0; Mismatches	2;	Indels	3; Gaps 1;
Qy	1	atragctcgctgggtgggtgcacccatcccttgggctctgttcctgggtgcgtgcgtccatcccgatcc	60		
Db	206	atragctcgctgggtgggtgcacccatcccttgggctctgttcctgggtgcgtgcgtccatcccgatcc	265		
Qy	61	caggctactctcccaacaactctactcttttagagagactgtcagaataaccacac	120		
Db	266	caaggctactctggccaaacgtactcttttagagagactgtcagaataaccacac	325		
Qy	121	aagagatctactcccggtcccgacaggccatcccccaggaggacaggagagatctc	180		
Db	326	aaggatctactcccggtcccgacaggccatcccccaggaggacaggagatctc	385		
Qy	181	atctgtcacaacaaacgttcggggcccggtgcaggccatccacatggatcatg	240		
Db	386	atctgtcacaacaaacgttcggggcccggtgcaggccatccacatggatctc	445		
Qy	241	acctggatgacgactggagagactctgtcggatgtggatcgatctggag	300		
Db	446	acctggatgacgactggagagactctgtcggatgtggatcgatctggag	505		
Qy	301	caaggccaccacgtctgtgtccatcgccggcagaacctggcgatctgggc--	357		
Db	506	caaggccaccacgtctgtgtccatcgccggcagaacctggcgatctggcagg	565		
Qy	358	agttatccctctccgggttccatgtcggatgtggatggaaagactaccac	417		
Db	566	agttatccctctccgggttccatgtcggatgtggatggaaagactaccac	625		
Qy	418	tacccctccggcgccgggttccatgtcggatgtggatggaaatqtc	477		
Db	626	tacccctacccggcgccgggttccatgtcggatgtggatggaaatqtc	685		
Qy	478	acgactacacacaaatgttggccacccaaacaaatgttggccatgttc	537		
Db	686	acgactacacacaaatgttggccacccaaacaaatgttggccatgttc	745		
Qy	538	tccggaaatgttcggggaaatttggagacgtcgccgttcttttcgtaa	597		
Db	746	tccggaaatgttcggggaaatttggagacgtcgccgttcttttcgtaa	805		
Qy	598	tatctccaaaggaaatgtggatggaaatgtggccatcaaaatgttgg	657		
Db	806	tatctccaaaggaaatgtggatggaaatgtggccatcaaaatgttgg	865		
Qy	658	gaggcccccccgatgggaaatgtggatggaaatgtggccatcaaaat	717		
Db	866	gaggcccccccgatgggaaatgtggatggaaatgtggccatcaaaat	925		
Qy	718	tacactccaaacccatggggaaatgtggatggaaatgtggccatccgg	777		
Db	926	tacactccaaacccatggggaaatgtggatggaaatgtggccatccgg	985		
Qy	778	gaaaccatgttttgtccaaacccggatgtggatggaaatgtggccatcc	837		
Db	986	gaaaccatgttttgtccaaacccggatgtggatggaaatgtggccatcc	1045		
Qy	838	ggggcaactacaaacggatgtggatggaaatgtggccatccggatgt	897		
Db	1046	ggggcaactacaaacggatgtggatggaaatgtggccatccggatgt	1105		
Qy	898	gggtccacgtgttggccaaacccggatgtggatggaaatgtggccatcc	957		
Db	1106	gggtccacgtgttggccaaacccggatgtggatggaaatgtggccatcc	1165		
Qy	958	tttggaaatgtgttcatggatggatggatggatggatggatggatgg	1017		
Db	1166	tttggaaatgtgttcatggatggatggatggatggatggatggatgg	1225		
RESULT 8					
ID	AAD17757	standard; DNA; 2412 BP.	XX		
AC	AAD17767;		XX		
DT	10-DEC-2001 (first entry)		XX		
DE	Human novel trypsin inhibitor-like protein, Nov-4e encoding DNA.		XX		
KW	Human; NOV-X Protein; KIAA1233-like protein; STE20-like protein; tumour; trypsin inhibitor-like protein; gene therapy; haematopoietic; illness; immunological disorder; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immunomodulatory; pharmacogenomic; haemostatic; human immunodeficiency virus; HIV; fertility disorder; neuroprotective; cytostatic; nontropic; anti-infertility; cancer; Nov-4e protein; ds. virus; Homo sapiens.		XX		
OS			XX		
FR	Key	Location/qualifiers			
FT	CDS	/tag= a			
FT		/product= "Human novel trypsin inhibitor-like protein, Nov-4e"			
FT	sig_peptide	206..271			
FT	mat_peptide	/tag= b			
FT		/tag= C			
FT		/product= "Human mature novel trypsin inhibitor-like protein, Nov-4e"			
PN	W020162928-A2.		XX		
PD	30-AUG-2001.		XX		
PF	26-FEB-2001; 2001WO-US06151.		XX		
PR	25-FEB-2000; 2000US-0184951.		XX		
PR	28-FEB-2000; 2000US-0185548.		XX		

PR 01-MAR-2000; 2000US-0185967.  
 PR 18-APR-2000; 2000US-0197723.  
 PR 27-APR-2000; 2000US-0199957.  
 PR 23-FEB-2001; 2001US-0789390.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 PI Vernet CAM, Fernandes E, Shimkets RA, Macdougall J, spaderna SK;  
 KX XX  
 DR WPI: 2001-582051-65.  
 DR P-PSDB; AAF10619.  
 XX  
 PT New isolated KIAA1233-like, STE20-like, or trypsin inhibitor-like polypeptide for diagnosing and treating pathological disorders, such as Parkinson's disease and for use in pharmacogenomics -  
 PT  
 XX  
 PS Claim 9; page 90-91; 189pp; English.  
 XX  
 CC The invention relates to novel human polypeptides referred to as NOV-X and their corresponding nucleic acid sequences. NOV-X collectively include NOV-1, NOV-2a and NOV-2b which are novel KIAA1233-like polypeptides, NOV-3a, NOV-3b, NOV-3c and NOV-3d which are novel STE20-like polypeptides and NOV-4a, NOV-4b, NOV-4c, NOV-4d and NOV-4e which are novel trypsin inhibitor-like polypeptides. NOV-X is used to identify a potential therapeutic agent that can modulate its activity and can be used for treating a pathology related to aberrant expression or aberrant physiological interactions of NOV-X. NOV-X or its DNA is used to determine the presence or predisposition to a disease associated with altered levels of NOV-X. NOV-X, its DNA and its antibody are used to treat or prevent a pathology associated with NOV-X. The pathologies states that can be treated or prevented are haematopoietic, cancer, immunological, tumour, neurodegenerative (e.g. Alzheimer's and Parkinson's disease), human immunodeficiency virus (HIV) illness and fertility disorders. NOV-X and its DNA are used in pharmacogenomics for predictive medicine. NOV-X DNA is used in gene therapy. The present sequence is a DNA encoding human novel trypsin inhibitor-like protein.  
 CC NOV-4e.  
 CC  
 CC Sequence 2412 BP; 560 A; 649 C; 688 G; 515 T; 0 other;  
 XX  
 SQ  
 Query Match 97.4%; Score 1452.6; DB 22; Length 2412;  
 Best Local Similarity 98.9%; Pred: No. 0;  
 Matches 1487; Conservative 0; Mismatches 4; Indels 12; Gaps 0  
 1 atggactgtggccgggtgtcatccccctggggctgtgttctgggtctggatcc 60  
 206 atggactgtggccgggtgtcatccccctggggctgtgttctgggtctggatcc 265  
 Db 61 caaggctacccctggcccaactgtactcttttagggactgtctggatccaaataccagcac 120  
 266 caaggctacccctggcccaactgtactcttttagggactgtctggatccaaataccagcac 325  
 Qy 121 aaccgatcttcactccgggtccggcggagccccatcccgaggagggacaaggagatcc 180  
 326 aacgagtgcttcactccgggtccggcggagccccatcccgaggagggacaaggagatcc 385  
 Db 181 atgtgcacacacaaatggggcggggcggggcggccatccggaggagggacaaggatcc 240  
 386 atctgcacacacaaatggggcggggcggggcggccatccggaggagggacaaggatcc 445  
 Qy 241 acctggatgacaactggagaactgtggatccatccggcggccatccggaggatcc 300  
 446 acctggatgacaactggggcggggcggccatccggaggatcc 505  
 Db 301 cacggccacccatgtcttctgggtccatccggcggccatccggaggatcc 360  
 506 cacggccacccatgtcttctgggtccatccggcggccatccggaggatcc 565  
 Qy 361 tacgcttcgggttcaatggggcggccatccggatggatggatggatggatcc 420  
 566 tacgcttcgggttcaatggggcggccatccggcggccatccggaggatcc 625











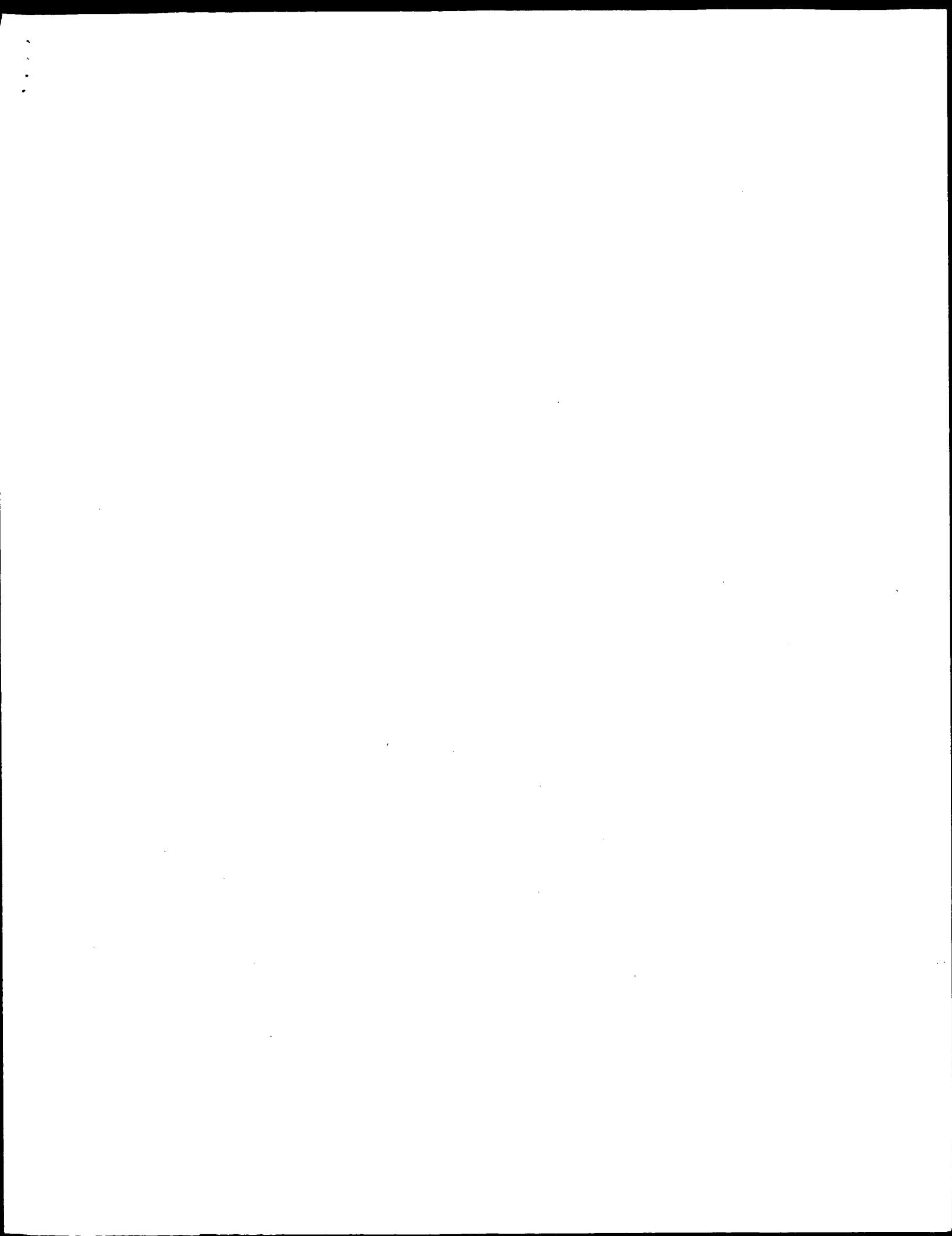






Tue May 28 16:10:01 2002

us-09-667-380a-1.rng



AC. NO: Q9HOB8  
Database: SPTREMBL-19

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

**OM protein - protein search, using sw model**

Run on: May 25, 2002, 14:27:40 ; Search time 122.23 Seconds  
(without alignments)

Title: US-09-667-380A-2  
Perfect score: 2759  
Sequence: 1 MSCVLLGVIPQIGLFLVCGS..... SESLGTPRDGKAFRIFAVFQ 497

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Maximum first A5 summaries

17	323.5	11.7	236	11	Q9DAG6
18	319.5	11.6	255	11	Q9CWG1
19	301.5	10.9	244	11	Q9X1K3
20	294.5	10.7	332	11	Q9CQ35
21	293.5	10.6	203	4	Q9H108
22	292.5	10.6	233	6	Q96L6
23	286.5	10.4	380	5	Q9VFY2
24	286.5	10.4	392	5	Q960R5
25	273.5	9.9	301	5	Q95QF6
26	273.5	9.9	220	5	Q16854
27	273.5	9.9	243	11	Q9R114
28	268.5	9.7	243	11	Q8B205
29	254.5	9.2	334	6	Q9BE36
30	240.5	8.7	165	6	Q77720
31	239.5	8.7	220	5	Q9T149
32	239.5	8.7	250	11	Q9D259
33	237.5	8.6	196	5	Q62507
34	235.5	8.5	220	5	Q4A932
35	234.5	8.5	424	5	Q76744
36	234.5	8.5	424	5	Q9XZ41
37	234.5	8.5	220	5	Q44331
38	234.5	8.5	249	6	Q9XD3
39	233.5	8.4	176	10	Q9J186
40	233.5	8.4	212	5	Q9BID5
41	232.5	8.4	212	5	Q96764
42	231.5	8.4	208	5	Q18543
43	231.5	8.4	217	6	Q77719
44	230.5	8.3	190	10	Q9SW05
45	230.5	8.3	212	5	Q20609

Database : SPTREMBL\_19;\*

Database : SPTREMBL\_19:\*

	RESULT	1
4:	sp-human:*	
5:	sp-invertebrate:*	
6:	sp-mammal:*	
7:	sp_mhc:*	
8:	sp-organelle:*	
9:	sp-phage:*	
10:	sp-plant:*	
11:	sp-rodent:*	
12:	sp-virus:*	
13:	sp_vertebrat:*	
14:	sp-unclassified:*	
15:	sp_xvirus:*	
16:	sp_bacteriap:*	
17:	sp_archeap:*	
		PRED. NO. IS THE NUMBER OF RESULTS PREDICTED BY CHANCE TO HAVE A SCORE GREATER THAN OR EQUAL TO THE SCORE OF THE RESULT BEING PRINTED, AND IS DERIVED BY ANALYSIS OF THE TOTAL SCORE DISTRIBUTION.
	SUMMARIES	
Result	Query	
No.	Score	Match Length DB ID Description
1	2756	99.9 497 4 Q9HB8B 09hb08 homo sapien
2	2071	75.1 371 4 Q96k61 homo sapien
3	1987.5	72.0 434 11 Q92cr3 mus musculu
4	1640.5	59.5 500 4 Q9hr336 homo sapien
5	1613	58.5 523 13 Q98s15 098s15 mus musculu
6	1536.5	55.7 507 11 Q99mm6 099mm6 rattus norv
7	968	35.1 188 11 Q92016 092016 gallus galli
8	758.5	27.5 258 13 Q98st6 098st6 gallus galli
9	740	26.8 258 11 Q99mm7 099mm7 mus musculu
10	739.5	25.8 4 Q43692 homo sapien
11	646.5	23.4 253 4 Q9H3Y0 0913y0 homo sapien
12	639	23.0 120 4 Q961b1 0961b1 hominidae
13	357	12.9 415 5 Q04228 0963k2 homin sapien
14	359	12.9 466 4 Q963k2 0963k2 homin sapien
15	349	12.6 489 11 Q9JJ56 091j56 mus musculu
16	349	12.6 489 11 Q9ET66 09et66 mus musculu
		PRELIMINARY; PRN; 497 AA..
ID	Q9HB8B	
AC	Q9HB8B;	
DT	01-MAR-2001 (TREMBLE, 16, Created)	
DT	01-MAR-2001 (TREMBLE, 16, Last sequence update)	
DT	01-DEC-2001 (TREMBLE, 19, Last annotation update)	
DE	HYPOTHETICAL	
GN	DKEZPA34Ba44.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostom	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=TESTIS;	
RA	Wambutt R., Heubner D., Mewes H.W., Gassnerbauer J., Wiemann S.; Submitted (Jul-2001) to the EMBL/GenBank/DDBJ databases.	
RL		
DR	EMBL: AL136061; CAB66795.1; -.	
DR	HSSP; P04284; 1CFE.	
DR	InterPro; IPR004043; LCCL.	
DR	InterPro; IPR01283; SCP.	
DR	PFam; PF0188; SCP; 1.	
DR	PRINTS; PRO0837; V5TPX1KE.	
DR	SMART; SM0138; SCP; 1.	
DR	FROSTIE; PS01010; SCP-AG5_PR1_SC7_2; 1.	
KW	HYPOTHETICAL PROTEIN.	
SQ	SEQUENCE 497 AA; 55919 MW; 1E7ADDCA37CA4ED CRC64;	
Query	Match Similarity	99.9%; Score 2756; DB 4; Length 497;
Best Local Matches	Conservative	99.8%; Pred. No. 3.5e-259; 0; Indels 0; Mismatches 496;
QY	1 MSCVIGGIVGIPGIIGLITVCGSGQYLLPNTVILEELLSKQHNEHSRSRRAIPREDKIV	
Db	1 MSCVIGGIVGIPGIIGLITVCGSGQYLLPNTVILEELLSKQHNEHSRSRRAIPREDKIV	
QY	61 MLHNKURGQVQPQASNNEYMTWDDELEKSAAWASOCWCWHEGPPLSLVSIQNLGA	



Query Match Score 1640.5; DB 4; Length 500;  
Best Local Similarity 59.5%; Pred. No. 9.5e-151;

Matches	285;	Conservative	73;	Mismatches	112;	Indels	20;	Gaps	6;
QY	13	LLFLVCGSGQYJLPPNTLELLSKYKHNSHH-----SRVRAALPREKEEILMLHNKL	66						
Db	14	LLFTIAQAVSAMVLPNATLLELKYMDEGEWVIANVSR-KRAITDSMDQSTILDLHNLK	72	:	:             :	:		:	
Qy	67	RGAQVOPASMEYMTWDELEKSAAAWASQCIWHEGPTSLVLVSIGQNLGAHNGRYRSPGF	126						
Db	73	RQCVYPPASNHEYMTWDELEKSAASWAEATCOLWHEGPASLPSIGONQAGHNGRYRPTF	132						
QY	127	HVOSWDEVKTYTYPSPSECNPWCERCSQPMCHYTOVWANTNKIGCAVNTRKMW	186						
Db	133	HVQAWDEVKDYTYPSPSECNPWCERCSQPMCHYTOVWANTNKIGCAVNTRKMW	192						
QY	187	GEWENAVYFCVNNSPKGMWIGEAPYKNGRCSPCOPSPSYGSGRNNTCYR--EETYTPK	243						
Db	193	QGIWPKAVYLVCNNSPKGNNWGHAPKCPCCACPPSFGGSCRENLYCRESERPYSPH	252						
QY	244	PETDENNEVE-TAPIPEENHVWLOPVRK-----TKPRKTSAYNTMOPVQRDTK	294						
Db	253	EPEEETNETERQRSKAQDATAOERPRTSPSSTGSDSEKEVISTQOMSQIVCEVL	312						
QY	295	KDRKGSTCNRYQCPAGCLNHAKIFGSLFYESSSSCTRAAHTHYGLDDKGGLDTRNG	354						
Db	313	RDOQKGTCNRQCPAGCLNHAKIFGSLFYESSSSCTRAAHTHYGLDDKGGLDTRNG	372						
QY	355	KVPEFKVKSERGVQSLSKYKPSKPKSSEMVSKVKVQDLCYTTAQLCPFEKATHCPRIHC	414						
Db	373	RKNYEIKSYRNGQSIKGQYQSANSTVSKVTQAVTCETTVEOLCPFPKAPSHCPVYC	432						
QY	415	AHKDEPSYWAQFGTNYIADISSICKTAVHAGTISNEGGDVTMPVKKTVGSLR	474						
Db	433	HNCMQANPHYARVIGTRIYSDISSICRAAVAHAGVYRNE-GGYVDVMPVDPDKRKYIASFLN	491						
QY	475	GVOSESGLGTP	484						
Db	492	GTFESQIOPQ	501						
RESULT	6								
Q9MM6									
ID	Q99MM6	PRELIMINARY;	PRT;	507 AA.					
AC	Q99MM6								
DT	01-JUN-2001	(TREMBREL. 17, Created)							
DT	01-DEC-2001	(TREMBREL. 17, Last sequence update)							
DE	COCACRISP.								
OS	Mus musculus (Mouse).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OX NCBI_TAXID=10090;								
RR	[1] SEQUENCE FROM N.A.								
RA	Smith D.M., Collins-Racie L.A., Lavallie E.R., Gamer L., Roberts D.J., RA Marigo V.A., Copeland N.G., Jenkins N.A., McCoy J., Tabin C.J.; RT CoccaCrisp, provides insight into the process of septation in the developing chicken midbrain;" submitted (DSC-2000) to the EMBL/GenBank/DBJ databases.								
DR	HSSL; P04284; ICFP; InterPro; IPR04043; LCCL; DR InterPro; IPR001283; SCP; DR Pfam; PF00188; SCP; DR PRINTS; PR00837; V5TPXLIKE. DR SMART; SM00198; SCP; DR PROSITE; PS01010; SCOP_AGS_5_PR1_SC7_2; DR SEQUENCE 507 AA; 57882 MW; 55408 DEE29A50A02 CRC64;								
Query Match	55.7%;	Score 1536.5;	DB 11;	Length 507;					
Best Local Similarity	55.4%;	Pred. No. 1.2e-140;							
Matches	274;	Conservative	74;	Mismatches	112;	Indels	35;	Gaps	7;
RESULT	7								
Q9ZU6									
ID	Q9ZU6	PRELIMINARY;	PRT;	188 AA.					
AC	Q9ZU6								
DT	01-MAY-1999	(TREMBREL. 10, Created)							
DT	01-MAY-1999	(TREMBREL. 10, Last sequence update)							
DE	01-JUN-2001	(TREMBREL. 17, Last annotation update)							
GN	LATE GESTATION LUNG PROTEIN 1.								
OS	Rattus norvegicus (Rat).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. RN NCBI_TAXID=10116;								
RP	[1] SEQUENCE FROM N.A.								
RC	STRAIN=NSTAR; TISSUE=LUNG; RX MEDLINE-9929450; Published=200362728;								
RA	Kaplan F., Ledoux P., Kassamali F.Q., Gagnon S., Post M., Koehler D., DeInning J., Sweezy N., "A novel developmentally regulated gene in lung mesenchyme: homology to a tumor derived trypsin inhibitor"; Am. J. Physiol. 276:L1027-L1036(1999).								
RT	EMBL; APL09674; ADI0986.1; HSSP; P04284; ICFP. DR InterPro; IPR001283; SCP. DR Pfam; PF00188; SCP; DR PRINTS; PR00837; V5TPXLIKE. DR SMART; SM00198; SCP; DR PROSITE; PS01010; SCOP_AGS_5_PR1_SC7_2; DR SEQUENCE 188 AA; 21907 MW; 5B6CC77515360BCE CRC64;								
Query Match	35.1%;	Score 968;	DB 11;	Length 188;					
Best Local Similarity	36.5%;	Pred. No. 5.9e-86;							
Matches	160;	Conservative	10;	Mismatches	15;	Indels	0;	Gaps	0;

QY	61	MLNKLRGQVQPOASMEYMTWDDLELELSAAWASOCIWHEGPTSLVSLVSGQNGLGAHGR 120	Db	23B	GSCTDNLCF 246
Db	1	MLNKLRLQRQVQPPASMEYMTWDDLEELSAAWASOCIWHEGPTSLVSLVSGQNGLAVHGR 60			
QY	121	YRSRPFHQWYDEVKDYPYPPSCNPNCPERCSPGMTHIQTWATTNKIGAVNC 180			
Db	61	RKMTVGEVENAVFVCNSPKGWNIGEAPYKNRGPCBCCPSYGGSCRNNLYCREWY 240			
QY	181	YRSRPFHQWYDEVKDYPYPPSCNPNCPERCSPGMTHIQTWATTNKIGAVNC 180			
Db	121	RSMSYWGTDITENAVYLVCNTSPKGWNIGEAPYKHPGRCPSCBCCPSYGGCRNNLYCREWY 240			
QY	241	TPKE 245			
Db	181	HQKPE 185			
	RESULT	8			
Q98ST6		PRELIMINARY; PRT; 258 AA.			
ID	Q98ST6				
AC	Q98ST6;				
DT	01-JUN-2001 (TREMBlrel. 17, Created)				
DT	01-JUN-2001 (TREMBlrel. 17, Last sequence update)				
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)				
DE	SUGARCRISP.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TAXID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21184121; PubMed=11287197;				
RA	Smith D.M., Collins-Racie L.A., Marigo V.A., Roberts D.J., Davis N.M.,				
RA	Hartmann C., Schweitzer R., Lavallie E.R., Gamer L., McCoy J.,				
RA	Tabin C.J.,				
RT	"Cloning and expression of a novel cysteine-rich secreted protein family member expressed in thyroid and pancreatic mesoderm within the chicken embryo."				
RT	Mech. Dev. 102:221-226(2001).				
DR	EMBL: AF329196; AAC16494.1; -.				
DR	HSSP; P04284; ICFE.				
DR	InterPro; IPR001283; SCP.				
DR	Pfam; PF00188; SCP; 1.				
DR	PRIMIS; PR00837; V5TPXLIKE.				
DR	SMART; SM00198; SCP; 1.				
DR	PROSITE; PS01010; SCP_AC5_PR1_SC7_2; 1.				
DR	SEQUENCE; 258 AA; 29150_MN; ZFA9949A2562640 CRC64;				
	Query Match	26.8%	Score	740	DB 11, Length 258;
	Best Local Similarity	55.7%	Pred. No.	1_3e-63	
	Matches	131	Conservative	31	Mismatches 61; Indels 12; Gaps
QY	13	LLFLFVGCGQGYLLPNVTLLELLSKYHNES-----HSRRAATPREDKEIL 60			
Db	12	LLSLICEATRVVLNPDLSSPLSPANNFIDTEALSTPLIESADIPKARKRYISQNDMIAIL 71			
QY	61	MLNKLRLQRQVQPOASMEYMTWDDLEELSAAWASOCIWHEGPTSLVSLVSGQNGLGAHGR 120			
Db	72	DYHNRQVRKFVPPDANMEYMVMDNLAKSAAWAATC1WDHGPSYLRLFLGSQNLSTFRGR 131			
QY	121	RKMTVGEVENAVFVCNSPKGWNIGEAPYKNRGPCBCCPSYGGSCRNNLYCREWY 240			
Db	132	YRSILOQVLPWDEVKDYPYPPSCNPNCPERCSPGMTHIQTWATTNKIGAVNC 180			
QY	181	YRSRPFHQWYDEVKDYPYPPSCNPNCPERCSPGMTHIQTWATTNKIGAVNC 180			
Db	192	QNNWWSWWRARWYLVCNTSPKGWNIGEAPYKHPGRCPMCYPMTHIQTWATTNKIGAVNC 235			
	RESULT	10			
QY	1	MSCVLLGGVLPGLFLVCGQSYLLPNVTLLELLSKYQ-----HNES---HSR 46			
Db	7	ISCVF-----LFSILCETSLAVLPLNSTDL--LLSNNNFTDELAALAHLSAKIPKAR 57			
QY	47	VRRALPREDKEEILMLNKLRLQRQVQPOASMEYMTWDDLEELSAAWASOCIWHEGPTSL 106			
Db	58	RKRVTSQNDMIALDHYNQVGRKVFPASMEYMTWDDLEELSAAWASOCIWHEGPTSL 106			
QY	107	LVSGTQNTGAHGWGRYRSPGFFVQWSWDEVKDYPYPPSECNPNCPERCSPGMTHIQTW 166			
Db	118	LRFGLQNLNSVTRGKRSIQLQVFWIDEVKDYPYPPDCNRCPCMRCYPMTHIQTW 177			
QY	167	WATPNKIGCAVNTRKMTVGEVENAVFVCNSPKGWNIGEAPYKNRGPCBCCPSYGGSCRNNLYCREWY 226			
Db	178	WATSNRIGCAIHTCQNMNVWGSVWRAYVLCNTSPKGWNIGEAPYKHPGRCPMCYPMTHIQTW 237			
	RESULT	2			
QY	227	GSRNRLCY 235			

RA	Kanamori T., Miyazaki K.
RT	"CDNA cloning of a novel trypsin inhibitor with similarity to human brain cancer proteins, and its frequent expression in human brain cancer cells".
RT	Biochim. Biophys. Acta 1395:202-208(1998).
RL	D45027; BAA206.1; -.
DR	HSSP: P04284; ICFE.
DR	InterPro: IPR01283; SCP.
DR	Pfam: PF00188; SCP; 1.
DR	PRINTS: PRO0837; VSTPXLKE.
DR	SMART: SM00198; SCP; 1.
DR	PROSITE: PS01010; SCP_AGE_PR1_SC7_2; 1.
SQ	SEQUENCE: 258 AA; 29065 MW; 1915A5831637795F CRC64;
RESULT 11	
ID	Q9H3Y0
AC	Q9H3Y0;
DT	01-MAR-2001 (TREMBLrel. 16, Created)
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT	01-MAR-2001 (TREMBLrel. 19, Last annotation update)
DR	EMBL; AL11382; CAC04190.1; -.
DR	HSSP: P04284; ICFE.
DR	InterPro: IPR01283; SCP.
DR	Pfam: PF00188; SCP; 1.
DR	PRINTS: PRO0837; VSTPXLKE.
SQ	SEQUENCE: 253 AA; 28605 MW; 28976C6F32E3887E CRC64;
Query Match	26.8%; Score 739.5; DB 4; Length 258;
Best Local Similarity	56.4%; Pred. No. 1.5e-63;
Matches	133; Conservative 30; Mismatches 60; Indels 13; Gaps 2;
QY	13 LLF-LVCGSGQYLLPNTLLELLSKYQHNE-----HSRVRAIPREDKEIL 59
Db	11 LIFSLIICEASTWVLNSTDSSSPPTNNFTDIBAALKAOQLSDADIPKARRKRYTSQNDMIAI 70
QY	60 IMLHNKLKGQVQPOASNMETMWDELEKSAAWASOQCINWEHGPTSLVLSIGQNGLAHWG 119
QY	[1] :   :
QY	118 WGRYRSPGFHVWSQWYDVEKDVTYPSECPNCWPGPERCSPGMCTHYTQWATNKIGCAV 177
Db	125 SGQYRSVYDLMWSWEEKWHLYFPARDNCNPHCPWCGPCSYHQMWNASSNRGCAI 184
QY	178 NTCKRMVWGEVWENAVYFVCMYSPKGWIGEAPYKNGRCPSCPPGSCRNRLCY 236
Db	185 HTCSSISVWGNWTHRAAYLVCNAYAIGNWIGESPYKMGKPCSSCPSPYQGSCNSNMCFK 243
RESULT 12	
ID	Q961B1
AC	Q961B1;
DT	01-DEC-2001 (TREMBLrel. 19, Created)
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE	HYPOTHETICAL 13.5 KDA PROTEIN.
OS	Homo sapiens (Human).
OC	Eukaryota; Metacida; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Cetartiodactyla; Homidae; Homo.
RA	TUSSUE-MELANOMA; STRAUBERG R.; Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL	InterPro: IPR01283; AAC01689.1; -.
DR	EMBL; BC001689; AAC01689.1; -.
Db	SEQUENCE FROM N.A.
QY	1 MSCVILGVIPIGLLEFGSGQYLLPNTLLELLSKYQHNE-----HSRVRAIPREDKEIL 60
Db	1 MSCVILGVIPIGLLEFGSGQYLLPNTLLELLSKYQHNE-----HSRVRAIPREDKEIL 60
QY	61 MHNKLKGQVQPOASNMETMWDELEKSAAWASOQCINWEHGPTSLVLSIGQNGLAHWG 120
Db	61 MHNKLKGQVQPOASNMETMWDELEKSAAWASOQCINWEHGPTSLVLSIGQNGLAHWG 120
RESULT 13	
ID	Q44228
AC	Q44228
DT	01-JUN-1998 (TREMBLrel. 06, Created)
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE	HRPT-1 PROTEIN.
GN	Halocynthia roretzi (Sea squirt).
OC	Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Stolidobranchia; Pyuridae; Halocynthia.
OX	NCBI_TaxID:7729;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=98267369; PubMed=9501991;
RA	Hotta K., Takahashi H., Satoch N;
RT	"Expression of an ascidian gene in the tip of the tail-bud stage embryos.";
RL	Dev. Genes Evol. 208:164-167(1998).
DR	EMBL; AB009609; BRA24011.1; -.
DR	HSRP; P04284; ICFE.
DR	InterPro; IPR001283; SCP.
DR	InterPro; IPR000884; TSP1.
DR	Pfam; PF00188; SCP; 1.
DR	PRINTS; PRO0837; VSTPXLKE.
DR	SMART; SM00198; SCP; 1.

DR SMART; SM00209; TSPI; 1.  
 DR PROSITE; PS0109; SCP\_AG5\_PRL\_SC7\_1; 1.  
 DR PROSITE; PS01010; SCP\_AG5\_PRL\_SC7\_2; 1.  
 DR PROSITE; PS50092; TSPI; 1.  
 SEQUENCE 415 AA; 47365 MW; 22366EB3F19DED3 CRC64;

Query Match 12.9%; Score 357; DB 5; Length 415;  
 Best Local Similarity 28.1%; Pred. No. 4.2e-26;  
 Matches 110; Conservative 58; Mismatches 128; Indels 96; Gaps 18;

QY 24 LIPNVVLEELLSKY-----QHNEHSRVRRA-----IPREKEET-----LM 61  
 Db 74 ILSEKEVVEETARNDIGLDPDANETFS-VKAVKAVK-TIPKIMDKIVLKPKFRMALLR 132  
 Qy 62 LHNKLKGQVOPOASNMWYMTWDDELEKSAAWASOCTIWEHGPT---SILVSIONGLAW 118  
 Db 133 THNARRAIAQPKAAWMRMMWDELERLAVAYSKCIEHNPKTHSREFYVBNLFT 192  
 Qy 119 GRYRSPGF--HVSWSYDEVKDYYTPYPSECNPWCPCERCS-GPKCTHYTQIWATNKIG 174  
 Db 193 GYAFTPSLKHAVERWDDEKQIYDE-----TKCQORGKMGHHTQVWADTFKMG 243  
 Qy 175 CAVNPKRKMYWGEWENAVYFCYNTSPKGWNWIGEAPYKNGRGCSECPSYGGSCRNNLC 234  
 Db 244 CGVIRCSDDIVGRGRWKNAIYLAVENYFGGNY-PTHPFTARPCKSKAPT--DCCRNLIC 300  
 Qy 235 YREEHYTPRPETDLMNEV--ETAPIPEENHVW-QPRVMRPTKRTSAVNWMQVVRCDT 292  
 Db 301 -----NNYTRDRKLDRKDIKWSEWTW--SSCSKSCGGVSTREROCNT 343  
 Qy 293 KMKURCKG---STCNRYQCPCAGLNLHAKITGSLFYESSSSICRAHYGSTDKGGL 347  
 Db 344 FVPGDCKDOPSEYVKFCVKPC-----KAEMEGN-----GGSFSYINVMNOG-- 384  
 Qy 348 VDTTRNGKUPFFVKSERHGVSISKYSKPSSSE 379  
 Db 385 -DKLKGSLQ-----QALQKHLGSPF 405

RESULT 14

ID Q959K2 PRELIMINARY; PRM; 266 AA.  
 AC Q959K2; PRELIMINARY; PRM; 266 AA.  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DE GLUTOMA PATHOGENESIS-RELATED PROTEIN (UNKNOWN) (PROTEIN FOR MGC:21203).  
 GN RTVE1.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1].

RP SEQUENCE FROM N.A.  
 RA Jang J.S., Hahn Y., Chung J.H.;  
 RT "Identification of novel mouse cysteine-rich protease inhibitor gene";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB04537; BAB0398; 1; -.  
 DR HSSP; P04244; JCPE.  
 DR InterPro; IPR01283; SCP.  
 DR Pfam; PF00188; SCP; 1.  
 DR PRINTS; PRO0837; V5MPXLIKE.  
 DR SMART; SM00198; SCP; 1.  
 DR PROSITE; PS0100; SCP\_AG5\_PRL\_SC7\_1; 1.  
 DR PROSITE; PS01010; SCP\_AG5\_PRL\_SC7\_2; 1.  
 DR PROTEASE; PS0100; SCP\_AG5\_PRL\_SC7\_1; 1.  
 DR PROTEASE; PS01010; SCP\_AG5\_PRL\_SC7\_2; 1.  
 SQ SEQUENCE 489 AA; 52676 MW; 01C207BE12E3CB9D CRC64;

Query Match 12.6%; Score 349; DB 11; Length 489;  
 Best Local Similarity 37.2%; Pred. No. 3.2e-25;  
 Matches 77; Conservative 28; Mismatches 86; Indels 16; Gaps 6;

QY 50 ALPREDEKEETILMLHNLKGQVOPOASNMWYMTWDDELEKSAAWASOCTIWEHGPTSLVS 109  
 Db 20 ALTEDEKQINWLHNLQYRAQVSPPASDMIQMRWDELAKAYAQOKCYWGINKER-GR 77  
 RP SEQUENCE FROM N.A.  
 RA Wu J., Zhang B., Zhou Y., Peng X., Yuan J., Qiang B.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2].

RP SEQUENCE FROM N.A.

RC TISSUE=RENAL ADENOCARCINOMA;

RA Strauberg R.; Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF000440; AAC02489; 1; -.  
 DR EMBL; BC012510; AAC12510; 1; -.  
 SQ SEQUENCE 266 AA; 30366 MW; C0466C0B5EFB2B7A CRC64;

Query Match 12.9%; Score 355; DB 4; Length 266;  
 Best Local Similarity 35.7%; Pred. No. 3.6e-26;  
 Matches 82; Conservative 33; Mismatches 75; Indels 40; Gaps 12;

Db 5 LATIARMWSFVSNISHT---ANILPDIEREDFIRDCVRHNRSEYKPTASDMLYMTWD 61  
 QY 84 DELERSAAWASQCCWHEG-----PTSL---LVSIGONGAHHRGKRYSPGFHQQS---W 131  
 Db 62 PALQIKAQKAWNSQOFSHNTRLPKPHKLPNFTSILGENI---W-TGSVPLFEVSSATNW 117  
 QY 132 YDEVKDYTYPYPSECNPWCPCERCSGPMCTHYTQIWATNKIGCAVNCTCRKMTVWGVWE 191  
 Db 118 YDEIQDYE-----KTRICKKVCGHYTQVWADSYKVGCAVFCPKVSGF-DALS 166  
 QY 192 NAVFVCNISPKSN-----WIGEAQYKNGPCSCRCPPSTGGSCRNLCYREE 238  
 Db 167 NGAHFICNYGGGNYPTW---PYKRGATCSACPNN - DKCLDNLCVNQR 210

RESULT 15

ID Q95J56 PRELIMINARY; PRM; 489 AA.  
 AC Q95J56; PRELIMINARY; PRM; 489 AA.  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DE CYSTEINE-RICH PROTEASE INHIBITOR CRIFI.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1].

RP SEQUENCE FROM N.A.  
 RA Jang J.S., Hahn Y., Chung J.H.;  
 RT "Identification of novel mouse cysteine-rich protease inhibitor gene";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB04537; BAB0398; 1; -.  
 DR HSSP; P04244; JCPE.  
 DR InterPro; IPR01283; SCP.  
 DR Pfam; PF00188; SCP; 1.  
 DR PRINTS; PRO0837; V5MPXLIKE.  
 DR SMART; SM00198; SCP; 1.  
 DR PROSITE; PS0100; SCP\_AG5\_PRL\_SC7\_1; 1.  
 DR PROSITE; PS01010; SCP\_AG5\_PRL\_SC7\_2; 1.  
 DR PROTEASE; PS0100; SCP\_AG5\_PRL\_SC7\_1; 1.  
 DR PROTEASE; PS01010; SCP\_AG5\_PRL\_SC7\_2; 1.  
 SQ SEQUENCE 489 AA; 52676 MW; 01C207BE12E3CB9D CRC64;

Query Match 12.6%; Score 349; DB 11; Length 489;  
 Best Local Similarity 37.2%; Pred. No. 3.2e-25;  
 Matches 77; Conservative 28; Mismatches 86; Indels 16; Gaps 6;

QY 50 ALPREDEKEETILMLHNLKGQVOPOASNMWYMTWDDELEKSAAWASOCTIWEHGPTSLVS 109  
 Db 20 ALTEDEKQINWLHNLQYRAQVSPPASDMIQMRWDELAKAYAQOKCYWGINKER-GR 77  
 RP SEQUENCE FROM N.A.  
 RA Wu J., Zhang B., Zhou Y., Peng X., Yuan J., Qiang B.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2].

RP SEQUENCE FROM N.A.

RC TISSUE=RENAL ADENOCARCINOMA;

RA Strauberg R.; Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF000440; AAC02489; 1; -.  
 DR EMBL; BC012510; AAC12510; 1; -.  
 SQ SEQUENCE 266 AA; 30366 MW; C0466C0B5EFB2B7A CRC64;

Query Match 12.9%; Score 355; DB 4; Length 266;  
 Best Local Similarity 35.7%; Pred. No. 3.6e-26;  
 Matches 82; Conservative 33; Mismatches 75; Indels 40; Gaps 12;

QY 25 LPNTVLEELLSKYOHNEHSRVRRAIPRED-KETILMLHNLKGQVOPOASNMWYMTWD 83

✓ P. 1-2

only 14 data in file  
Result 14 is entry 32 cont. on page 1  
in file 40M (32 cont. on page 1)

Copyright (c) 1993 - 2000 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on:

May 25, 2002, 10:30:04 ; Search time 2776.47 seconds

(without alignments)  
1237.819 Million cell updates/sec

**Title:** US-09-667-380A-1  
**Perfect score:** 1491  
**Sequence:** 1 atggatctgcgttcgggtgg.....ggatctttgtgtcaggac 1491

**Scoring table:** IDENTITY\_NUC  
**Gapop** 10.0 , Gapext 1.0

**Searched:**

1797656 seqs, 10463268293 residues

**Post-processing:** Minimum Match 0%  
**Maximum Match** 100%  
**Listing first 45 summaries**

**Database :**

1: GenBank:  
 2: gb\_btgc:  
 3: gb\_in:  
 4: gb\_com:  
 5: gb\_ov:  
 6: gb\_pat:  
 7: gb\_ph:  
 8: gb\_pl:  
 9: gb\_pr:  
 10: gb\_ro:  
 11: gb\_sts:  
 12: gb\_sy:  
 13: gb\_un:  
 14: gb\_v:  
 15: em\_ba:  
 16: em\_fun:  
 17: em\_hum:  
 18: em\_in:  
 19: em\_mu:  
 20: em\_on:  
 21: em\_or:  
 22: em\_ov:  
 23: em\_pat:  
 24: em\_ph:  
 25: em\_pl:  
 26: em\_ro:  
 27: em\_sts:  
 28: em\_un:  
 29: em\_v:  
 30: em\_htg\_hum:  
 31: em\_btgc\_inv:  
 32: em\_htgc\_other:  
 33: em\_htgo\_inv:  
 SUMMARIES

## ALIGNMENTS

**RESULT**

1

AX101173

AX101173

LOCUS

Sequence 1 from Patent WO0121651.

DEFINITION

1491 bp

ACCESSION

W0121651

VERSION

AX101173.1 GI:13620006

KEYWORDS

human.

ORGANISM

Homo sapiens

REFERENCE

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS

(bases 1 to 1491)  
Donoho,G., Turner,C.A., Wattler,F., Neils,M., Friedrich,G.,

TITLE

Novel human protease inhibitor-like proteins and polynucleotides

encoding the same

and is derived by analysis of the total score distribution.

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1491	368	a	BASE COUNT

FEATURES

Lexicon Genetics Incorporated (US)

Source

Location/Qualifiers

1. /organism="Homo sapiens"

/cb\_xref="txon:9606"

1. .1491

ORIGIN

Db	961	GGAAGTCCTTCATGAAAGCTGTCTACATATGCCGCCCATCCACTACGGATC	1020
Qy	1021	crtgatgacaaaggagggcttggatattaccaggaaaggaaagtcccttttgtg	1080
Db	1021	CTGGATGACAAAGGGAGGCCTGGTATACACCAGGAAGGGAGTCCCTTCGG	1080
Qy	1081	aactctggagacacggcggtgcgtccatcagaataccaaaccttcacgttcat	1140
Db	1081	AACTCTGAGAGACACGGCGTGCAGTCCCCTACGAAATTCAACCTTCAGC	1140
Qy	1141	gttcacaaatgttgcaggatgttgcactgtcacacgtttgtcaatgttgc	1200
Db	1141	GTCATTAAGGTGAAGTCAGGAGTTGCACTTCACAGGACCTTGCACTG	1200
Qy	1201	tttggaaaaggccacgaaacttcacttgcggcaacttgcgtccggcacatgt	1260
Db	1201	TTGCAAAGACCCAGCACTACTGCCCCAAGATACTCCATGTCGGCACACT	1260
Qy	1261	ccttccttcgtggctccgtgttggaccaacatctatggatgttgcacatgc	1320
Db	1261	CCTTCCTACTGGCTCCGGTGTGAAACCAACTCATCTATGCAGATAC	1320
Qy	1321	aaacacagctgtgtacgcggggggatcatcgcaagagatgggggtgtacgtat	1380
Db	1321	AAGACAGCTGTGCAACGGGGAGTCATCGCAACGAGAGTGGGGTGACT	1380
Qy	1381	cccggtggaaaaaaaggaaacactacgtgggtcgttcacggatggatggatgt	1440
Db	1381	CCGGGGAAAGAACGACTTACGTGGCTCGTCAGGAATGGAGTCAGTCTGA	1440
Qy	1441	ctggggactcttggatggaaaggccctccggatctttgtcttcgtatggatgg	1491
Db	1441	CTGGGGACTCTCGGATGGAAAGGCCTTCGGATCTTGTGTCAGGCAG	1491

QY	121 aacqaqtctactcccccgggtccgcacagccccaggaggacaaggaggatcctc 180	Db	1416 TTTGAAAGCCACACTCACTGCCAGATTCATTCCTCCCACTGAGCGAA 1475
Db	336 AACGAGTCCTACTCCCGGTCCGCAGGCCATCCCAGGAGCACAGGAGATCTC 395	QY	1261 CCTTCTACTGGCTCGGTGTTGGRACCACTATGAGATCTCAAGCACTGC 1320
QY	181 atgtgcacacaagttggggccgggtggacttcggccatcgccacatggatcatg 240	Db	1476 CCTCTACTGGCTCCGGTGTGACCAACATCATGAGATCTCAAGCACTGC 1535
Db	396 ATGCTGACACAAAGCCTGGGGCCAGTGCACTGCACTGAGTCATGAGTACATG 455	QY	1321 aagcagatgtcaacggggatcatggacaaaggatggggactatgtgacg 1380
QY	241 acctggatggactggaaactggaaatctgtgtggcggcgtggcggatgtgg 300	Db	1536 AACAGCTGTCACGCCAGGAGTCATCGAACAGGAAGTGGGGTACGGTGATG 1595
Db	456 ACCTGGATGACGAGACTGAGAGTCGTCAGCGAGTCAGTCATGCACTGGAG 515	QY	1381 cccgtggataaaaaggaaagacacttgcgtggctcaacttgcgtggatc 1440
QY	301 cacggggccaccatgtgtgtgtgtccattggcggacaaactggggcctactgg 360	Db	1596 CCCGGATAAAGATAACCTAGTGGGCTCCCTCAGAAGGAGTCAGTGAAGC 1655
Db	516 CAGGGCCACCAGTCGGGTCAGTCAGCTGGCAG 575	QY	1441 ctgggactctcggatggaaaggccctccggatcttgctgtggatcggatc 1491
QY	361 tatgcgttccgggttccatgtgcgtctgtatggactacacatc 420	Db	1656 CTGGGACTCTCGGATGGAAAGGCCCTCCGGATCTTGCTGCTGAGGCAG 1706
Db	576 TATGGCTTCCGGGTTCAGTCAGCTGGCAGTCACCTAC 635		
QY	421 ccctacccggagcggatgtgcacccctgggtccaggaggatggatgtgacg 480	RESULT 3	AX086850
Db	636 CCCTACCCGGAGTGCACCCCTGGGTTCAGAGACTGTCCTGGGGCT 695	LOCUS	AX086850
QY	481 caetacacacacatgttgtggccacccacaaacaaatgtcggtgtgtggacac 540	DEFINITION	Sequence 802 from Patent WO0112659.
Db	696 CACTACACACAGATAGTTGGCCACCAACACAGATGGTGTGCTGACACCCTC 755	ACCESSION	AX086850.1 GI:13276048
QY	541 cggaaatggactgtcgggggatgtggggacacggcggttactttgtctcaatt 600	VERSION	
Db	755 CGGAGATAGTGTCTGGGAGAGTGTGGGAGCT 815	KEYWORDS	
QY	601 tctccaaacgggaactgtggatggaaaggccctacaagaatggccggctgtctg 660	SOURCE	human.
Db	815 TCTCCAAGGGAACTGGATTGGAGAACGCCCTACAAACATGGCCGGCCTGCGCTGAG 875	ORGANISM	Homo sapiens
QY	661 tgccccaccaggatgtggatggcggatgtcgacaaacttgttacccggaaatc 720	REFERENCE	1 (bases 1 to 4574)
Db	876 TGGCCACCCAGCTATGGAGCAGTCAGAACACTTGTTACCGAGAAACTAC 935	AUTHORS	Wiemann,S.
QY	721 actccaaacacctgtggacggatgtggatgtggatggggatcccttcgtggaa 780	TITLE	Human dna sequences
Db	936 ACTCCAAACCTGTGACGAGTCAGTGGAGGAGACGGCTCCATCTGGAGA 995	JOURNAL	Patent: WO 0112659-A 802 22-FEB-2001;
QY	781 aaccatgtttgtgtccaaacggggatggatggacccaaaggaaacctctcg 840	FEATURES	German Human Genome Project (DE)
Db	996 AACCATGTTGGTGCACACCGAGGGTGTAGACCCACCAAGGAAACCTCTGC 1055	BASE COUNT	Location/Qualifiers
QY	841 gtcaactacatgtggatggatggatgtggatggatggatggatggatggatgg 900	source	1. 4574
Db	1056 GTCACACTCATGACCAAGTGTGACGAGTCAGACGTGCAAGGG 1115	/db xref="Homo sapiens"	/
QY	901 tccaaatggatggatggatggatggatggatggatggatggatggatggatgg 960	/db xref="taxon:9606"	
Db	1116 TCCACCTGTACAGGIFACACTGCCCCAGCAGCTGCTGACCAATACAGAC 1175	ORIGIN	
QY	961 ggaatgtgttctatgtaaatgtcgatgtatggccggccatccactacggatc 1020	Query Match	99.9%; Score 1487; DB 6; Length 4574;
Db	1176 GGAGCTGCTGTTATGAAAGCTGCTGATGCAATATGCCGCGCCACACTACGGATC 1235	Best Local Similarity	99.9%; Pred. No. 0; Mismatches 0; Gaps 0;
QY	1021 ctggatggatggatggatggatggatggatggatggatggatggatggatggatgg 1080	Matches	1489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db	1236 CTGGATGACAGGGAGGCCCTGGATATCACCAGAACCGAAAGTCCCTCTGCG 1295	QY	121 aacggatgtcaatccgggtggatggatggatggatggatggatggatggatgg 180
QY	1081 aatgtgttccatgtggatggatggatggatggatggatggatggatggatggatgg 1140	Db	325 AACGAGTCCTACTCCGGTGGCAGGCCATCCAGGGAGCACAGGAGAACCTC 384
Db	1296 ATGCTGAGAGCACGGGCGTGGAGTCCTACCAACTCAACTCTGGATCTG 1355	QY	181 atgtgcacacaacaaatggatggatggatggatggatggatggatggatggatgg 240
Db	1141 ggtgtcaaaatggatggatggatggatggatggatggatggatggatggatggatgg 1200	Db	385 ATGCTGACACAAAGCTGGGGCCAGTCAGGCTCAGGAGTCATGAGTCATG 444
QY	1356 GGTGCAAAAGTGAAGTCAGCGAGTGGACTGACTGCTACAGGACGGTGTGCG 1415	QY	241 acctggatggactggatggatggatggatggatggatggatggatggatggatgg 300
QY	1201 ttgtggaaaggccggcaactactggccaaatggatggatggatggatggatgg 1260	Db	505 CAGGGCCACCGTGTCTGGTCACTGGGAGAACCTGGGCTACGGGCTACGGCAGG 564
QY		QY	361 tatgcgttccgggttccatgtggatggatggatggatggatggatggatggatgg 420







REFERENCE	Fukuyama; Melarza; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrini; Hominidae, Homo.
AUTHORS	1 (sites)
TITLE	Lillie,J., Brown,J.L., Bolt,A. and van Huffel,C. Novel genes, compositions and methods for the identification, assessment, prevention, and therapy of human cancers
JOURNAL	Patent: WO 0179556-A8 884 25-OCT-2001;
FEATURES	Millennium Predictive Medicine, Inc. (US)
source	/organism="Homo sapiens" 1..477 location/Qualifiers
BASE COUNT	1183 a /db_xref="taxon:9606" 1201 c 1235 g 1257 t
ORIGIN	
Query Match	99.8%; Score 1487.8; DB 6; Length 4877;
best local similarity	99.9%; Pre. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches	1489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	1 atagactcgctctgggtgttcatcccttgggtgtgtttcttgttcgtatcgatcc 60
Db	4657 ATAGCTGGTCCGGGGCATCCCTGGGTGCTTCCTGGTGTCTGGCGATCC 4598
QY	61 caaggtaacctcgccaaacgtcaactctttagagagtgctcgcaataccggac 120
Db	4597 CAAGCTTACCTCCGCCAACGTCACTCTAGAGCTCTAGCAATACACAC 4538
QY	121 aacgatctcaactccgggtccacagaccatcccgccatccactaaggatcc 180
Db	4537 AACAGTCACICCGGTCAGGACATCCAGCTAGGATC 4478
QY	1021 ctggatgacaaaggggggccgggttataccggaaacccggaaagggtcccttcgg 1080
Db	3637 CTTGGATCACAAAGGGAGGCCCTGGGATAATACCGAACCGGAAGSTCCCTCTCGTG 3578
QY	1081 aagtctgagacacgcgggtcagtcataccaaaccttcccaagtcattcaag 1140
Db	3577 AAGTCTGAGACACCGCGTCACTGCAGTCCTCAAACTTCACACTCATCG 3518
QY	1141 gtgtcaaaaggaaaggatggactgtcttcacacgcgtgtctgtggcc 1200
Db	3517 GTGTCAAAGTGAAGGCCAGGATGGCTACACGCCGTTGCGCTGGCCCG 3458
QY	1201 ttggaaaaggccgcaactcaacttgcggaaagaaatccatgtccggcacactcaaaccgaa 1260
Db	3457 TTGGAAAGGCCAACCTACCTGCCTGGCCAAAGAACATTCATGTCGGCACACTCAAAGCGAA 3398
QY	1261 ctttcctactgggtcggtttggaaaccacatcatggagatracctaaagctctgc 1320
Db	3397 CCTCTCTACTGGCTCCGGTGTGAGACAAACATCATGTCAGATACCTAGCATCTGC 3338
QY	1321 aagacagatgttgcacggggatcatcagaacaaagggtggggacgtcgatgtgatc 1380
Db	3337 AAGCACGCCGTGACCGGGGACTCATCAGCAACGAGGTGGGGTACGGCGATG 3278
QY	1381 ccgtgtataaaaaggaaacctacgggggtcgatggatggatctcaatccaaac 1440
Db	3277 CCCGTGGAAAAGAACGACCTACGTGGCGRCGCTCAGGATGGAGTCAGTGTGAAAGC 3218
QY	1441 ctggggactctcggtatggatggatggatccatgttgcgtcgcc 1491
Db	3217 CTGGGGACTCTCGGGATGGAAAGGCCCTCCGGATCTTGGCTGTCAGCGAG 3167
RESULT	7
AX285079/_C	AX285079 4877 bp DNA linear PAT 20-NOV-2001
DEFINITION	Sequence 884 from Patent: WO0179556.
ACCESSION	AX285079
VERSION	AX285079.1 GI:17045767
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens

















Tue May 28 16:10:00 2002

us-09-667-380a-1.rge

b. (A), (S)  
Page 16

Qy	1163	atttggactcttacacggccgttctcaudtgttcccdttgtaaaaggccaaactact.	1222
Db	1284	CTGTGACTCTTGAAACACACTGTGGACAGCCTCISGCCATTCTAAGCTGCCTCACATT	1343
Qy	1223	gcceagaaccatgtccgcacactgtggaaagcggacccttctactgggtcccggt	1282
Db	1344	GCCCCAAGAGTATACttCTCGTAATCTGTCAGCAATCCACATATTGCTCGTGAA	1403
Qy	1283	ttggaccacatctatggataacttcggatctggaaagacagtgtggcccgccggag	1342
Db	1404	TGGAATCTCGAGTTTATCTGATCTGTCCTGATACTGCGAGAGCAGACTACANGCTGGAG	1463
Qy	1343	tcatcggcaagagtggggttacgtggacgttatggcccggtgataaaaggaaacct	1402
Db	1464	TGGTGTGAATNTCA---CGGGTTTGTGATGPAATGCTGAGACAAAGAACCT	1520
Qy	1403	acgtggggctctcggaaatggacttcggatctaaaggccgtggactcttcggatgtggaa	1462
Db	1521	ACATGGCTCTTTCGAAATGGAATCTCTCAGAAAGTTACAGAATCCTCCAGGAGAA	1580
Qy	1463	aggccctccggatctttgcgt	1484.
Db	1581	AGGCATTCAGAGTGTGCT	1602

Search completed: May 25, 2002, 13:22:02  
Job time: 10318 sec

Tue May 28 16:10:02 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: May 25, 2002, 10:25:14 ; Search time 2064.9 Seconds  
(without alignments)  
9745.731 Million cell updates/sec

Title: US-09-667-380A-1  
Perfect score: 1491  
Sequence: 1 attagatctgcgtccctggggg.....ggatcttttgtgtcaggag 1491

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Egn:  
1: em\_estba: \*  
2: em\_eastnum: \*  
3: em\_eatin: \*  
4: em\_eatmu: \*  
5: em\_estrov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_htc: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_htc: \*  
12: gb\_gss: \*  
13: em\_gss\_hum: \*  
14: em\_gss\_inv: \*  
15: em\_gss\_pln: \*  
16: em\_gss\_vrt: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result NO.	Score	Query Length	DB ID	Description
1	967	64.9	2329 11 AK019034	AK019034 Mus muscu
c	2	509	34.1 651 10 BF361052	BF361052 MR2-OP007
c	3	447.2	30.0 482 10 BG011536	BG011536 PM1-SN018
c	4	439.8	29.5 459 10 BFF98318	BF998318 QV0-GN021
c	5	434.2	29.1 900 10 BG175425	BG175425 602337780
c	6	421.4	28.3 680 10 BII11002	BII11002 603382445
c	7	411.8	27.6 748 10 BI55378	BI55378 603382445
c	8	401.9	26.9 835 10 BFI18143	BFI18143 6018565
c	9	399.2	26.8 580 9 AW910974	AW910974 ur81F08Y
c	10	398.6	26.7 762 10 BI15480	BI15480 602903664
c	11	395	26.5 406 10 BII0828	BII0828 115-MTU226
c	12	391.2	26.2 822 10 BFI1923	BFI1923 601790112
c	13	387.8	26.0 783 10 BI683391	BI683391 603315836
c	14	383.6	25.7 410 10 BFR753946	BFR753946 QV2-CT055
c	15	379.2	25.4 163 10 BG861553	BG861553 60795324
c	16	358.8	25.2 91 10 BI1248527	BI1248527 60292944
c	17	358.8	24.1 573 10 BB310264	BB310264 601089361

## ALIGNMENTS





sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 2020663  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL:  
<http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t2=PM1-GN01811200-004-d10&t3=2000-12-11&t4=1>  
 Seq primer: puc 18 forward  
 High quality sequence start: 14  
 High quality sequence stop: 482.  
 Location/Qualifiers  
 1. 482  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone.lib="GN0180"  
 /dev\_stage="Adult"  
 /note="Organ: Placenta, normal; Vector: puc18; Site.1: Sm;  
 ; Site.2: Smal; A mini-library was made by cloning  
 products derived from ORESTES PCR (U.S. Letters Patent  
 application No. 1,167,716 - Ludwig Institute for Cancer  
 Research) profiles into the PUC 18 vector. Reverse  
 transcription of tissue mRNA and cDNA amplification were  
 performed under low stringency conditions."  
 136 a 127 c 126 g 93 t  
 h  
 Similarity 30.0%; Score 447.2; DB 10; Length 482;  
 55; Conservative 0; Pred. No. 1.6e-103; Mismatches 13; Indels 0; Gaps 0  
 ggccggaaaggatgactgtctggggagaagtttggggaaacgcgggtctacttttgtctgcaa 596  
 GCTCGGGATGAGCTGCTGTTAGAAGTTRGGAGAACGGCCTACTATGTCGCAA 74  
 atttcggccaaaggaaactgtatggggagaacccctatacaagaatggccggccgtgtc 656  
 ATTCTCCAAAGGGGAACTGGATGGAGAACCCCTACACAGATGGCGGCCCTGCRC 134  
 atggcccaaccatgtggcgatgtggggaaactgtgttccggaaaggaaac 716  
 ACTGCCCAACCCTACTATGGGGCGCTGGAGAACACCTGGAGAAC 194  
 actccaaacctgtggaaacggccggatgtggggaaacgggtccatctgtgg 776  
 ACACCTCCAAACACTGAAACGGGAGAGATGAGGTGGAACCGGCTCCCATTCCTGA 254  
 aaaccatgttgctccaaagggtatggggaaaccccaaggccaaaccc 836  
 ggtccacgtgtggccaaatgtggccaaaggatgtggatggggaaatgtgg 956  
 GGTCGACTATAATGACCCAAAGTGCTGAGCACCRAGATGAGGGAGGCCA 374  
 gttggatgtggccaaatgtggccaaaggatgtggatggggaaatgtgg 314  
 ggatgtggccaaatgtggccaaaggatgtggatggggaaatgtgg 896  
 GTCCACGTTGAGACAGGTACCAAGTGCTGAGCACCRAGATGAGGGAGGCCA 434  
 tggaaagtctgttctatgtggaaatgtggccatataatccggccccc 1004







Query Match 26.8%; Score 399.2; DB 9; Length 580;  
Best Local Similarity 80.5%; Pred. No. 3-1e-91; Mismatches 113; Indels 0; Gaps 0;

FEATURES source  
Match 1 .762  
Organism="Mus musculus"  
/db\_xref="taxon:10990"  
/clone="IMAGE:5033703"  
/clone\_1db="NIR\_CGP\_Man3"  
/tissue\_type="tumor, gross tissue"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: PCMV-SPORE6; Site:1: NotI; Site:2: Sall; Cloned unidirectionally. Primer: Oligo dT. Average Insert 2 kb. Library constructed by Life Technologies catalog #12017-1018. Investigators providing samples: Lothar Hennighausen/Chu-xia Deng, NIH Reference samples: Lothar Hennighausen/Chu-xia Deng, NIH Reference (1999). Note: this is a NCICBP library."  
BASE COUNT 164 a 234 c 217 g 147 t  
ORIGIN

---

Query Match 26.7%; Score 398.6; DB 10; Length 762;  
Best Local Similarity 78.1%; Pred. No. 5e-91; Mismatches 134; Indels 0; Gaps 0;

FEATURES source  
Match 1 .762  
Organism="Mus musculus"  
/db\_xref="taxon:10990"  
/clone="IMAGE:5033703"  
/clone\_1db="NIR\_CGP\_Man3"  
/tissue\_type="tumor, gross tissue"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: PCMV-SPORE6; Site:1: NotI; Site:2: Sall; Cloned unidirectionally. Primer: Oligo dT. Average Insert 2 kb. Library constructed by Life Technologies catalog #12017-1018. Investigators providing samples: Lothar Hennighausen/Chu-xia Deng, NIH Reference samples: Lothar Hennighausen/Chu-xia Deng, NIH Reference (1999). Note: this is a NCICBP library."  
BASE COUNT 164 a 234 c 217 g 147 t  
ORIGIN

---

Query Match 26.7%; Score 398.6; DB 10; Length 762;  
Best Local Similarity 78.1%; Pred. No. 5e-91; Mismatches 134; Indels 0; Gaps 0;

FEATURES source  
Match 1 .762  
Organism="Mus musculus"  
/db\_xref="taxon:10990"  
/clone="IMAGE:5033703"  
/clone\_1db="NIR\_CGP\_Man3"  
/tissue\_type="tumor, gross tissue"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: PCMV-SPORE6; Site:1: NotI; Site:2: Sall; Cloned unidirectionally. Primer: Oligo dT. Average Insert 2 kb. Library constructed by Life Technologies catalog #12017-1018. Investigators providing samples: Lothar Hennighausen/Chu-xia Deng, NIH Reference samples: Lothar Hennighausen/Chu-xia Deng, NIH Reference (1999). Note: this is a NCICBP library."  
BASE COUNT 164 a 234 c 217 g 147 t  
ORIGIN

---

RESULT 10

BI155480  
LOCUS BI155480 DEFINITION 5' mRNA sequence. ACCESSION BI155480 VERSION BI155480.1 SOURCE EST. KEYWORDS house mouse.  
ORGANISM Mus musculus. REFERENCES 1. (bases 1 to 762) National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999). Contact: Robert Strausberg, Ph.D. Email: cgps@remail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Chu-xia Deng Ph.D. DNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLM at: http://image.llnl.gov Plate: LILM1093 row: o column: 16

---

RESULT 11

BI030828/C LOCUS BI030828 mRNA LINEAR EST 14-JUN-2001

Tue May 28 16:10:02 2002

**DEFINITION** IL5-MT0265-290301-421-a05 MT0265 Homo sapiens cDNA, mRNA sequence.  
**ACCESSION** B1030828  
**VERSION** B1030828.1 GI:14437458  
**EST.**  
**KEYWORDS**  
**SOURCE**  
**ORGANISM** Homo sapiens  
**EUKARYOTA;** Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
**MAMMALIA;** Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
**REFERENCE** 1 (bases 1 to 406)  
Dias Neio,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Navea,M.A., da Silva,W. Jr., Zago,M.A., Bordim,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.P., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., de Oliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
**TITLE** shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
**JOURNAL** proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
**MEDLINE** 20020663  
**COMMENT** Ludwig Institute for Cancer Research  
Laboratory of Cancer Genetics  
Rua Prof. Antonio Prudente  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?ti=IL5t2-IL5-MT026>)  
290301+421-a05st=2001-03-29&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 354.  
Location/Qualifiers  
1. . 406  
Organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_11b="MT0265"  
/dev\_stage="Adult"  
/note="Organ: marrow; Vector: puc18; Site\_1: SmaI; Site\_2:  
note="A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No.  
,716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."  
85 a 112 c 105 g 104 t  
**FEATURES**  
**SOURCE**  
**BASE COUNT**  
**ORIGIN**  
Query Match 26.5%; Score 395; DB 10; Length 406;  
Best Local Similarity 99.8%; Pred. No. 3.1e-90; 0; Indels 1; Gaps  
Matches 406; Conservative 0; Mismatches 0;  
OY 1076 tcgtgaagtcgagaacacaqqcggtgcaggctccatggcaatacaacttcacgtcat 1135  
Db 406 TCGGAAGTCGAGAGCACGGCGCCAGTCCTCAGAACATAC-AACCTCCAGCAT 348  
OY 1136 tcatgggtgtcaaaatggaaatgcggatgttgactgtcacacggcgttgccatgt 1195  
Db 347 TCATGGGTGTCAAAGTGAAGTCGAGGATTTCGACTACAGACGCCAGTCAGCT 288  
OY 1196 gccgtttgaaaaggccgcaactcaactggccaaagaattccatgttgcggccacatggcaag 1255  
Db 287 GCCCTTGAAAGCCAGCACTCACTGCCAGATTCATGTCGGCACACTCCAAAG 228  
OY 1256 acgaaccttcactgggtccgtgttggaccacatctatggatataactcaaga 1315  
Db 227 AGGAACCTTCATGGGTCGGGTTGGACACAGATCATPATGCGATACCTCAGCA 168  
OY 1316 tctgcggacacgtgtgcacggggagtatcggacacggatgtggggtaatggacg 1375  
Db 167 TCTGCAGACAGCTGAGCAGCGGGAGTCATCAGCACAGAGATGGGGTGAAGTCAG 108

**Sex:** "Female; **tissue\_type:** "virgin"  
**dev\_stage:** "5 months";  
**lab\_host:** "DH10B";  
**note:** "Organ: mammary; Vector: pCAG-SV40T6; Site\_1: sAL1;  
Site\_2: NotI; Cloned unidirectionally; Primer: Oligo dr.  
library constructed by Life Technologies. Investigator  
providing samples: Jeffrey Green, M.D., NIH"  
244 C 223 g 149 t

```

/sex="female", "virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DHL0B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SALL
Library constructed by Life Technologies. Primer: Oligo dT
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT          167 a      244 C      223 g      149 t
ORIGIN             244 C      223 g      223 g      149 t

Query Match          Best Local Similarity      26.0%; Score 387.8; DB 10; Length 783;
Matches            492; Conservative 0; Pred. No. 2, 9e-88; Mismatches 137; Indexes 2; Gaps 2
QY
1 atggcgctggctcgtggatgttcatcccttggggctgtgttcgttcgtccggatcc 60
Db 155 ATGAGCTGTCCTGTCGAACATATGGCTCTGAGGGCCTTGCTCTGCTGTGAGTA 214
QY
61 caaggctactctggcccaatgttccttttagggactgtcgcaagaataccggac 120
Db 215 CAAGGCCATTGTTCTCCACACACCAGAGCTGGAGACTGCTGAGAAATTCCACAT 274
QY
121 aacgagtttcactcccggggttcgcggacatcccaaggagacaaggagagatctc 180
Db 275 GCAGAGCCACACTCACGGCTGCAGGCCATCCCATGTCAGACCCCCAGAGATCTC 334
QY
181 atggccatccacaaatctccggccacagtgcggccatccatggatcatgtggat 300
Db 335 ATGCTACACACAAAGCTGGCCAGGGTCAACATGGTCAACATGGAGACATG 394
QY
241 acctggatgacggactctggatggatgttcgtggatgtggatgtggatgtggat 454
Db 395 ACTGGGGATGGAGGAGTGGAGAATGAGACATTCAGCGGTGGCACATGAGCTGGGG 514
QY
301 cacggccaccaatcgatgttcgtggatccacatggatcatgtggatcatgtggat 420
Db 455 CACGGGCCGGCCGGCCCTCTGGTCATCCACGGGAGAACCTGGTGGTGGGGAGG 574
QY
361 tacgtctcccggttccatgtggatcttgtatgtggatgtggatgtggatgtggat 360
Db 515 TACCGCCTCTCCGGGTCATGTCAGTGGAGGAGGATTAACCTAC 634
QY
421 ccttggatggatgttcgtggccatccatggatgtggatgtggatgtggatgtggat 480
Db 575 CCATACCCCAACGAGTCCTCCACGGGTCGGGGTGTCAACGCCCATGTCAC 694
QY
481 cactacacaaatgtttggccacccaaacatgtggatgtggatgtggatgtggat 540
Db 635 CACTACACAGTCGGGACCTACACAAACAGATGGCTGGCCACACCTGGGG 753
QY
541 cggagatgtactgtctggggaaaatgtggaaaggcggttctacttttgtctgtcaatrat 600
Db 695 CGGACACGAACTCTGGGAGACACTGGGAATGCC-GTAACTCTCTGCAATTAT 753
QY
601 tctccaaaggggaaatgtggatgtggaaatcc 631
Db 754 TCCTCC-CAGGGAAACCTGGATGGAGGGCC 783

RESULT          14
BF753946
LOCUS          BF753946
DEFINITION      QV3-CT0556-041000-370-f05 CT0556 mRNA linear EST 10-JAN-2001
ACCESSION      BF753946
VERSION         BF753946.1
SOURCE          GI:12080622
KEYWORDS        EST.
ORGANISM        human
ORGANISM        Homo sapiens
ORGANISM        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE       1 (bases 1 to 410)

```



Qy 301 cacyggccaccagt-ctgtcggtccatcgggagaacctggcgcgtcaactgggacag 359  
Db 449 CACGGCCGCCGCGCCTTCAGAGGRCATGGGAGAACCTGGCTTGCACTGGGCAG 508  
Qy 360 gtagtcgtctccgggggtgggtccatgtcgtagtgcctcggtatgacgggttaaggactaaccta 419  
Db 509 GTACCGCTCCTGCGTCCATGTGCACTCATGGTACGACGAGGTGAGGARTACACCTA 568  
Qy 420 cccttaccccgagcggtgcacccctgggtcgccaggaggtctcggtccatggcac 479  
Db 569 CCCATPACCCCCACGAGTGCACTCCAGGTGCCGGGAGCGGGCTCAAGCCCCATGTCAC 628  
Qy 480 gcaactacacacagatgttggccaccacaacaaatcggtgtgtgaaacctg 539  
Db 629 CCACTRACACAGATGTCCTGGCCACCA-CAACRAAGATCGGCTGRCGCGRCACACTG 687  
Qy 540 ccggaaaatgtactgtcggtggagaaatggagaaacgcgcgtctactttgtgtcaatta 599  
Db 688 TCGGAACATGAACTGGAGACTTGAGATGCCATGTCGCTGCAATT 747  
Qy 600 ttctccaaggggac 615  
Db 748 TTCTCCCCAAGGGAAC 763

Search completed: May 25, 2002, 12:35:04  
Job time: 7790 sec

GenCore version 4.5  
copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2002, 10:33:09 ; Search time 70.13 Seconds

(without alignments)  
5222.295 Million cell updates/sec

Title: US-09-667-380a-1  
perfect score: 1491  
Sequence: 1atgatgtgcgtccctgggtgg.....ggatcttgcgtcaaggcag 1491

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Maximum DB seq length: 0

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Issued\_Patents\_NA:  
1: /cgn2\_6/podata/2/lna/5B\_COMB.seq:\*\*  
2: /cgn2\_6/podata/2/lna/6A\_COMB.seq:\*\*  
3: /cgn2\_6/podata/2/lna/6B\_COMB.seq:\*\*  
4: /cgn2\_6/podata/2/lna/PC1US\_COMB.seq:\*\*  
5: /cgn2\_6/podata/2/lna/backfiles1.seq:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	74.6	5.0	970	US-08-773-368-2
2	74.6	5.0	970	US-09-199-887-2
3	43.2	2.9	7218	US-08-232-463-14
4	42.2	2.8	11219	US-07-642-34C-1
5	38.4	2.6	696	US-08-439-009A-1
6	38.4	2.6	30001	US-08-998-16-689
7	38.4	2.6	30001	US-08-125-468-1
8	37.4	2.5	1031	US-08-174-933-1
9	37.4	2.5	1211	US-08-937-362-40
10	37	2.5	1211	US-08-937-362-40
11	37	2.5	1211	US-08-973-970-40
12	37	2.5	1211	US-08-973-970-40
13	37	2.5	1211	US-09-035-855-40
14	37	2.5	1211	US-08-705-342A-40
15	37	2.5	1211	US-09-324-442-40
16	36.8	2.5	756	US-09-173-014-1
17	36.4	2.4	3468	US-07-951-71A-2
18	36.4	2.4	3468	US-08-459-448A-2
19	36.4	2.4	3468	US-08-459-595A-2
20	36.4	2.4	3468	US-08-455-504B-2
21	36.4	2.4	3468	US-08-455-542-40
22	36.4	2.4	3468	US-09-053-549-3
23	36.4	2.4	3468	US-03-547-422-2
24	35.8	2.4	1410	PCT-US95-15428-14
25	35.8	2.4	1817	US-08-564-496C-14
26	35.8	2.4	696	US-08-998-416-050
27	35.6	2.4		

RESULT 1  
US-08-773-368-2  
Sequence 2, Application US/08773368  
Patent No. 5856130  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Golik, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN PATHOGENESIS-RELATED PROTEIN  
TITLE OF INVENTION: PROTEIN  
NUMBER OF INVENTIONS: 1  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
ZIP: 94304  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ FOR Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/773, 368  
FILING DATE: 2002-05-21  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy RJ  
NAME: Billings, Lucy RJ  
REGISTRATION NUMBER: 316, 749  
REFERENCE/DOCKET NUMBER: PP-0186 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-555-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 970 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLogy: linear  
IMMEDIATE SOURCE:  
CLONE: 1599164  
SEQUENCE:  
US-08-773-368-2

Query Match Best Local Similarity 5.0%; Score 74.6; DB 2; Length 970;  
Pred. No. 3.9e-11;

Matches 275; Conservative 0; Mismatches 226; Indels 36; Gaps 3;

QY 146 gagccatccccaggagacaaggagatccatgtcacacaacaaagcttcggggc 205  
Db 88 GAGCCCTCACAGTGAATGGAGAACAGTTGATGTGGAGCTGCACAACTCPACCGGCC 147

QY 206 aggttcagccatgggtccacatggatgtactgtggatgtacgtcaactgtgagaat 265  
Db 148 AGGTATCCCCACGGCCTCACACTGTCACATGAGATGGAGACTGGCC 207

QY 266 ctgcgtcgatgtggccatgtggatgtacgtcatggatgtggatgtggatgt 325  
Db 208 TCGCCAGGGCTACGACGGCACTNNHGTNG-----GCCAACAGGAGCCGGC 261

QY 326 ccatcgggagaacttggggctactggggcaatgtatgtatcggtttccatgtc 385  
Db 262 GCGCGGGGAGAATCTGTCGCCATCACAGACGGCATGGTGCCCTGGCATGG 321

QY 386 agtcgtgtatggatgtggatgtggatgtggatgtggatgtggatgtggatgt 445  
Db 322 AGGTGTGGCACACGAGGCTGAGCACTACACCTAGCCCCGCACCT----- 369

QY 446 gggtggccatggatgtggatgtggatgtggatgtggatgtggatgtggatgt 505  
Db 370 -----ggatgtggatgtggatgtggatgtggatgtggatgtggatgtggatgt 417

QY 506 ccaccacaatggatgtggatgtggatgtggatgtggatgtggatgtggatgt 565  
Db 418 AGACAGAGGATGGCTGGTGTGTTCCACRTGAG-----CTCCAGGTGTTG 471

QY 566 ttggggatgtggatgtggatgtggatgtggatgtggatgtggatgtggatgt 625  
Db 472 AGGAGAACACATCGAACTACTGGTGTGCAACTATGAGCCTCGGGAACTGGAA 531

QY 626 aagcccccataaaggatgtggccgtgtcgatgtggatgtggatgtggatgt 682  
Db 532 AACGGCCCTACAGGGACTCCGTCGTCACAGTCCTCTGCTACTCGA 588

RESULT 2

US-09-199-887-2

; Sequence 2, Application US/09199887

; Patent No. 6071874

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; TITLE OF INVENTION: NOVEL HUMAN PATHOGENESIS-RELATED PROTEIN

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ FOR Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/199, 887

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/773, 368

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy RJ

REGISTRATION NUMBER: 36, 749

REFERENCE/DOCKET NUMBER: PR-0186 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 970 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: 1599164

US-09-199-887-2

Query Match 5 0%; Score 74.6; DB 3; Length 970;

Matches 275; Conservative 0; Mismatches 226; Indels 36; Gaps 3;

QY 146 gagccatccccaggagacaaggatgtggatgtggatgtggatgtggatgt 205  
Db 88 GAGCCCTCACAGTGAATGGAGAACAGTTGATGTGGAGCTGCACAACTCPACCGGCC 147

QY 206 aggttcagccatgggtccacatggatgtactgtggatgtacgtcaactgtgagaat 265  
Db 148 AGGTATCCCCACGGCCTCACACTGTCACATGAGATGGAGACTGGCC 207

QY 266 ctgtcgatgtggatgtggatgtggatgtggatgtggatgtggatgtggatgt 325  
Db 208 TCGCCAGGGCTACGACGGCACTNNHGTNG-----GCCAACAGGAGCCGGC 261

QY 326 ccatcgggagaacttggggctactggggcaatgtatgtatcggtttccatgtc 385  
Db 262 GCGCGGGGAGAATCTGTCGCCATCACAGACGGCATGGTGCCCTGGCATGG 321

QY 386 agtcgtgtatgtggatgtggatgtggatgtggatgtggatgtggatgtggatgt 445  
Db 322 AGGTGTGGCACACGAGGCTGAGCACTACACCTAGCCCCGCACCT----- 369

QY 446 gggtggccatggatgtggatgtggatgtggatgtggatgtggatgtggatgt 505  
Db 370 -----ggatgtggatgtggatgtggatgtggatgtggatgtggatgtggatgt 417

QY 506 ccaccacaatggatgtggatgtggatgtggatgtggatgtggatgtggatgt 565  
Db 418 AGACAGAGGATGGCTGGTGTGTTCCACRTGAG-----CTCCAGGTGTTG 471

QY 566 ttggggatgtggatgtggatgtggatgtggatgtggatgtggatgtggatgt 625  
Db 472 AGGAGAACACATCGAACTACTGGTGTGCAACTATGAGCCTCGGGAACTGGAA 531

QY 626 aagcccccataaaggatgtggccgtgtcgatgtggatgtggatgtggatgt 682  
Db 532 AACGGCCCTACAGGGACTCCGTCGTCACAGTCCTCTGCTACTCGA 588

RESULT 3

US-09-199-887-2

; Sequence 3, Application US/09199887

; Patent No. 6071874

; GENERAL INFORMATION:

; APPLICANT: DORNER, F.

; APPLICANT: SCHEFFLINGER, F.

; TITLE OF INVENTION: RECOMBINANT FOWIPOX VIRUS

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22213-0299

COMPUTER READABLE FORM:



OTHER INFORMATION: /function= "approximate span of acyltransferase 2 domain of module 1"

FEATURE: NAME/KEY: misc\_feature

LOCATION: 5574..6125

OTHER INFORMATION: /function= "approximate span of beta-ketoreductase domain of module 1"

FEATURE: NAME/KEY: misc\_feature

LOCATION: 6369..6626

OTHER INFORMATION: /function= "approximate span of beta-ketoreductase domain 2 of module 1"

FEATURE: NAME/KEY: misc\_feature

LOCATION: 6678..11219

OTHER INFORMATION: /function= "approximate span of acyl carrier domain 2 of module 1"

FEATURE: NAME/KEY: misc\_feature

LOCATION: 8262..9305

OTHER INFORMATION: /function= "approximate span of acyltransferase domain of module 2"

FEATURE: NAME/KEY: misc\_feature

LOCATION: 9906..10454

OTHER INFORMATION: /function= "approximate span of beta-ketoreductase domain of module 2"

FEATURE: NAME/KEY: misc\_feature

LOCATION: 10707..10964

OTHER INFORMATION: /function= "approximate span of acyl carrier domain of module 2"

US-07-642-734-C-1

---

Query Match Best Local Similarity 2.8%; Score 42.2; DB 1; Length 11219; Matches 98; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 199 cggggccaaatgtggatcggccatccaaatggatcatgaccgtggatgaccaatgg 258  
Db 8324 CGGCAACACTCCGGCGCCATCCTTCCATGGGACCCGGCCAGGAAGACCAAC 8265

QY 259 gggaaatgtgtggatcgatggccatcgatgtggatggatggatggatggatgg 318  
Db 8264 TGAACCACCGGGGACACCCGGTGTACACGGGAAACACGGGACGCCGGCTGCAC 8205

QY 319 ctgggttcatcgatggatggatggatggatggatggatggatggatggatggatgg 378  
Db 8204 CCCCGCAACCGGTGTCACGGCCCTCCGGGTCGGCCACGGCCGAGCAC 8145

QY 379 catgtgcgtc 389  
Db 8144 GAAGTCTGGC 8134

RESULT 5  
US-07-439-009A-1/C  
; Sequence 1, Application US/08439009A  
; Patent No. 600487  
GENERAL INFORMATION:  
APPLICANT: Donadio, S  
APPLICANT: Katz, L  
APPLICANT: Mcalpine, J B  
TITLE OF INVENTION: Method of Directing Biosynthesis of Specific Polyretides  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:

---

ADDRESSEE: Steven F. Weinstock  
STREET: Abbott Laboratories D377/AP6D-2 One Abbott  
STREET: Park Rd  
CITY: Abbott Park  
STATE: IL  
COUNTRY: US  
ZIP: 60064-3500

COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/4439, 009A  
FILING DATE: 11-MAY-1995  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Casuto, Dianne  
REGISTRATION NUMBER: 40, 943  
REFERENCE/DOCKET NUMBER: 4952.US.DI  
TELECOMMUNICATION INFORMATION:  
PHONE: 847.938.3137  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11219 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Saccharopolyspora erythraea  
STRAIN: NRRL 2338

FEATURE: NAME/KEY: misc\_feature  
LOCATION: 744..6659  
OTHER INFORMATION: /function= "approximate span of module 1"  
FEATURE: NAME/KEY: CDS  
LOCATION: 744..11219  
OTHER INFORMATION: /function= "gene= "eryA""  
OTHER INFORMATION: /product= "'ORFL encoding modules 1 & 2 for 6-deoxyerythronolide B'"  
FEATURE: NAME/KEY: misc\_feature  
LOCATION: 744..1868  
OTHER INFORMATION: /function= "approximate span of acyltransferase domain 1 of module 1"  
FEATURE: NAME/KEY: misc\_feature  
LOCATION: 1998..2198  
OTHER INFORMATION: /function= "approximate span of acyl carrier domain 1 of module 1"  
FEATURE: NAME/KEY: misc\_feature  
LOCATION: 2220..3626  
OTHER INFORMATION: /function= "approximate span of beta-ketoreductase domain/module 1"  
FEATURE: NAME/KEY: misc\_feature  
LOCATION: 3831..4811  
OTHER INFORMATION: /function= "approximate span of acyltransferase 2 domain of module 1"  
FEATURE: NAME/KEY: misc\_feature  
LOCATION: 5574..6125  
OTHER INFORMATION: /function= "approximate span of beta-ketoreductase domain of module 1"  
FEATURE: NAME/KEY: misc\_feature

us-09-667-380a-1.rni

Tue May 28 16:10:01 2002

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATOR: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/998,416

FILING DATE: 24-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: CH 0016/97

FILING DATE: 31-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

INFORMATION FOR SEQ ID NO: 689:

SEQUENCE CHARACTERISTICS:

REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 689:

SEQUENCE CHARACTERISTICS:

REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 689:

SEQUENCE CHARACTERISTICS:

REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 689:

SEQUENCE CHARACTERISTICS:

REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 689:

SEQUENCE CHARACTERISTICS:

REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 689:

SEQUENCE CHARACTERISTICS:

REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 689:

SEQUENCE CHARACTERISTICS:

REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 689:

SEQUENCE CHARACTERISTICS:

REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

LOCATION: 6369..6626 /function= "approximate span of  
OTHER INFORMATION: acyl carrier domain 2 of module 1"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 6678..11219 /function= "approximate span of  
OTHER INFORMATION: /function= "approximate span of  
OTHER INFORMATION: module 2"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 6678..8066 /function= "approximate span of  
OTHER INFORMATION: /function= "approximate span of  
OTHER INFORMATION: beta-ketoacyl acpsynthase of  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 9906..10454 /function= "approximate span of  
OTHER INFORMATION: /function= "approximate span of  
OTHER INFORMATION: beta-ketoreductase domain of module 2"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 10707..10964 /function= "approximate span of  
OTHER INFORMATION: /function= "approximate span of  
OTHER INFORMATION: acyl carrier domain of module 2"  
US-08-439-009A-1

RESULT 6 998-416-689  
; Sequence 689, Application US/00998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippsen, Peter  
; APPLICANT: Rohmann, Rainer  
; APPLICANT: Steinke, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Juergen  
; APPLICANT: Knochle, Philipp  
; APPLICANT: Reblischung, Corinne  
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHYA GOSSYPII  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 1152  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 623964 Artis Corporation  
; STREET: 3024 Cornwallis Road  
; CITY: Research Triangle Park  
; STATE: No. 27709-380a-1  
; COUNTRY: USA

RESULT 7 998-416-689-1/c  
US-08-125-468-1/c  
; Sequence 1, Application US/08125468  
; Patent No. 559385  
; GENERAL INFORMATION:  
; APPLICANT: Ryan, Michael J.  
; APPLICANT: Latvin, Jason A.  
; APPLICANT: Strathy, Nancy  
; APPLICANT: Fantini, Susan E.  
; TITLE OF INVENTION: Cloning of the biosynthetic pathway for  
; chlortetracycline and tetracycline formation and cosmids  
; TITLE OF INVENTION: useful therein  
; TITLE OF INVENTION: useful therein  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: American Cyanamid Company  
; STREET: One Cyanamid Plaza  
; CITY: Wayne  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07470



US-08-997-080-40 Application US/08997080  
Sequence 40, Application US/08997080  
Patient No. 5960524  
**GENERAL INFORMATION:**  
APPLICANT: WATSON, JAMES D.  
TITLE OF INVENTION: METHODS AND  
NUMBER OF SEQUENCES: 194  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Ann  
STREET: 2601 Elliott Avenue, S  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
COMPUTER TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/997  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sleath, Janet  
REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 1100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-679-0555  
TELEFAX: 206-269-0563  
TELEX:  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1211 base Pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA

Query Match 2.5%; Score 37; DB 2; Length 1211;  
 Best Local Similarity 44.6%; pred. No. 0.84;  
 Matches 145; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-08-997-362-40

	Best Local Similarity	%	DB	Length
Matches	145;	Pred.	No.	121;
Qy	119	aaacacgttcactccgggtcccgagccatccaggagacaaggagatcc	0;	Mismatches
Db	383	ACACGGACTGTGACCCCCCCTTAAAGGGCCGACCTTGACCTAACATGGCA	178	Indels
				Gaps
				0;

TELEPHONE: 206-269-0565  
 TELEX: 206-269-0563  
 INFORMATION FOR SEQ ID NO: 400  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1211 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: Genomic DNA  
 US-08-873-970-40

Query	Match	2.5%	Score	37;	DB	3;	Length	1211;
Best	Local	Similarity		Pred.	No.	0	84;	Matches
Matches	145;	Conservative				0;	Mismatches	180;
QY	119	acacagactctactcccggtccggcagagcatcccaaggagaa	gacaggagatcc			0;	Indels	0; Gaps
Db	383	ACACCGACTGAGTCAGCCGCCCGTAAACAAAGGCCGACGGTGACTCAAGTGGGA	442			0;		
QY	179	tatgtgcataacaagaatgttggggcaggatgcggctcaaggctccacaatggatca	238			0;		
Db	443	CCCTCCGTGACCCAGAGCTCCGGGTGGTGTGAGGACCCGCTCAGCCGATCG	502			0;		
QY	239	tgcacctggatgacaaactggagaatgtgtggagtggtggcaggatgtatcggtgg	502			0;		
Db	503	GCAGCCGCCCTGCTGGCTGTGAGGGCGGTGGCGCGCGGAACTGGGACCTGGGAC	562			0;		
QY	299	agcacggccaccatgtgtgtcatcggtggacacttgtggcqctactaaatcgaa	350			0;		

-08-873-970-40  
Sequence 40, Application US/08873970  
Patent No. 6001361

**GENERAL INFORMATION:**  
**APPLICANT:** Tan, Paul

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR  
NUMBER OF SEQUENCES: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS  
CORRESPONDENCE ADDRESS: 106

ADDRESSEE: Law Offices of Ann W. Speckman  
STREET: 2601 Elliott Avenue Suite 1100

ZIP: 98121  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS  
SOFTWARE: FastSEO for Win

CURRENT APPLICATION DATA: 11/20/05 VEN

FILING DATE: 1-27-00  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/705 347

FILING DATE: 29-AUG-1996  
ATTORNEY/AGENT INFORMATION:

NAME: Sleath, Janet  
REGISTRATION NUMBER: 37-007

REFERENCE/DOCKET NUMBER: 11000.1002C1  
TELECOMMUNICATION INFORMATION:

APPLICATION NUMBER: 08/873,970  
FILING DATE: 12-JUN-1997  
APPLICATION NUMBER: 08/997,362  
FILING DATE: 23-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Sleath, Janet  
REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 11000.1002c3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-269-0565  
TELEX: 206-269-0563  
TELEFAX: 206-269-0563  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1211 base Pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-09-095-855-40

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08-705,347A  
FILING DATE: 28-AUG-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sleath, Janet  
REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 11000.1002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206.269.0565  
TELEFAX: 205.269.0563  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1211 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-705-347A-40

QUERY Match 2.5%; Score 37; DB 4; Length 1211;  
 Best Local Similarity 44.6%; Pre: No. 0.84; Mismatches 180; Indels 0  
 Matches 145; Conservative 145; Pairs 145; Score 37; DB 4; Length 1211;  
 Oy 119 aacaaggagtctcactccagggtcgccagggatccaggaggaaaggaggaa  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 383 ACACCGACTGTAACCGCCGCCGCTAACAGGGCCGACCGTGACCTTACAGTC  
 Qy 179 tcatactgcacaaacaagaatctcgaggccaaatggcggccatcgaggctccacatgaa  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 443 CCTTCTGACCCAGGCTTCCGGCTCTGGCAGGCCAACCGCCGGTAAAGCC  
 Qy 239 tgaccctggatgacaaactggaaatggatctgtcgacggcgccggcaatgtatgaa  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 503 GCGGCCCCCTCTGGCTGTGAGCTGGCGGTTTGGCCGGCTGACCTTGACCG  
 Qy 299 aacacggggccaccggcttcgttgttccatcgggaaacctggcgccactgg  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 563 ACCGGGAGGATCACTACAGGGCGGCGAATGCGCTCTGGACCCCTCCGAG  
 Qy 359 ggatccgtctccgggggtttccatgtggatctggatggatggatggatggatgg  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 623 GGTCGCCGTCCTGATCAGCACATCTGATGGGTGACGCCGGCTTCACGCCGAG  
 Qy 419 accccatccacccggccggatggatggatggatggatggatggatggatggatgg  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 683 TGTGGGGCAGACGGGGATCCC 707

**RESULT 14**  
**US-08-705-347A-40**  
 ; Sequence 40, Application US/08705347A  
 ; PATENT NO. 6,884,255  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tan, Paul  
 ; APPLICANT: Hirama, Jun  
 ; APPLICANT: Visser, Elizabeth  
 ; APPLICANT: Skinner, Marjorie  
 ; APPLICANT: Scott, Linda  
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND  
 NUMBER OF SEQUENCES: 55  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Speckman Picard PLLC  
 STREET: 2601 Elliott Avenue, Suite 4185  
 CITY: Seattle  
 STATE: WA  
 COUNTRY: USA  
 ZIP: 98121  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk

```

Query Match          Score: 37; DB 4; Length 1211;
Best Local Similarity 44.6%; Pred. No. 0; 84; Mismatches 0;
Matches 145; Conservative 0; Mismatches 180; Indels 0; Gaps 0;
QY 119 accacgacttcactccgggtcgacagccatcccgaggagacaaggagatcc 178
Db 363 ACACGACGAGCTGGCGCCCGCCGCGTACAGGGCCGACCTGACCTNAGTGGAGA 442
QY 179 tcatgtgcacaaacaaagtctcgggcagggtcgaccgtccaggccaaatggatca 238
Db 443 CCTTCCTGACCCAGAGCCTCCGGCGTACTCTGCAGGCCAACCGGGCGTAAGCCGACCG 502
QY 239 tgacctggatgacgaaactggagaatgtgtcgagctggggcagttgtatctgg 298
Db 503 GCAGGGCCGCTGGGGCTGGCTGAGGGCGGTGGCCCGCGCTGAACCTGGGACACGCC 562
QY 299 agcacggggccaccatgtcgatgtgtgtccatggcaggacccggcgatctggggca 358
Db 563 ACCGGAGAGTTCACTAGCGGGCTGAGTGGCGGTCTGACCCCTCCAGGGCT 622
QY 359 ggtatcgcttcgcgggttccatgtggatgtacggatgtggatgtggatgtggatgt 418
Db 623 GGCGCCGTCCTGATCACATCTCGATGATGGTGACGCCGGCGTCAGGCCGACAGACA 682
QY 419 acccttacccatggggatgggtgcaccc 443
Db 683 TGTGGGGAAAGACCGAGGGATCCC 707

RESULT 15
US-09-3-24-542-40
; Sequence 40, Application US/09324542
; Patient No. 6328978
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestridge, Ross
; APPLICANT: Prestridge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000_107C1
; CURRENT APPLICATION NUMBER: US/09/24, 542
; CURRENT FILING DATE: 1999-06-02
; EARLIER APPLICATION NUMBER: US 08/997, 080
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 1211
; TYPE: DNA
;
```

; ORGANISM: *Mycobacterium vaccae*  
US-09-324-542-40

Search completed: May 25, 2002, 13:23:33  
Job time: 10224 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

### OM protein - protein search, using sw model

Run on: May 25, 2002, 13:23:34 ; search time 61 seconds  
(without alignments)  
199.009 Million cell updates/sec

Title: US-09-667-380a-2  
Perfect score: 2759  
Sequence: 1 MSCVILGGVIVPLGLLFLVCGS.....SESLGTPRDKAFRIEAVRQ 497

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/podata/2/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/podata/2/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/podata/2/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/podata/2/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/podata/2/1aa/pCTUS.COMB.pep: \*  
6: /cgn2\_6/podata/2/1aa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No. Score Query Match Length DB ID

Description

Result No.	Score	Query Match	Length	DB	ID	Description
1	355	12.9	266	2	US-08-773-368-4	Sequence 4, Appli
2	355	12.9	266	3	US-09-199-887-4	Sequence 4, Appli
3	352	12.8	219	2	US-08-773-368-3	Sequence 3, Appli
4	352	12.8	219	3	US-09-199-887-3	Sequence 3, Appli
5	322	11.7	270	2	US-08-773-368-1	Sequence 1, Appli
6	322	11.7	270	3	US-09-199-887-1	Sequence 1, Appli
7	273	9.9	220	5	PCT-US96-07709-30	Sequence 30, Appli
8	273	9.9	248	5	PCT-US96-07709-25	Sequence 25, Appli
9	269	9.7	204	5	PCT-US96-07709-33	Sequence 33, Appli
10	256	9.3	204	1	US-08-419-414-9	Sequence 9, Appli
11	255	9.3	204	1	US-08-614-935-2	Sequence 2, Appli
12	256	9.3	204	3	US-09-130-287-2	Sequence 12, Appli
13	244.5	8.9	440	1	US-07-930-686-12	Sequence 1, Appli
14	244.5	8.9	440	2	US-08-460-998-12	Sequence 1, Appli
15	240	8.7	204	1	US-08-614-935-1	Sequence 1, Appli
16	240	8.7	204	1	US-08-614-935-1	Sequence 1, Appli
17	237	8.6	203	1	US-08-419-414-8	Sequence 8, Appli
18	237	8.6	203	1	US-08-614-935-3	Sequence 3, Appli
19	237	8.6	203	3	US-09-130-287-3	Sequence 2, Appli
20	236.5	8.6	424	1	US-08-419-414-2	Sequence 2, Appli
21	227	8.2	151	1	US-08-614-935-28	Sequence 28, Appli
22	227	8.2	151	3	US-09-130-287-28	Sequence 5, Appli
23	226	8.2	205	1	US-08-450-944-5	Sequence 5, Appli
24	226	8.2	205	1	PCT-US96-07709-5	Sequence 5, Appli
25	226	8.2	221	1	US-08-450-944-2	Sequence 2, Appli
26	226	8.2	221	5	PCT-US96-07709-2	Sequence 110, Appli
27	225.5	8.2	137	2	US-07-857-224B-110	Query Match

RESULT 1  
US-08-773-368-4  
; Sequence 4, Application US/08773368  
; Patent No. 5856130  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Goi, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN PATHOGENESIS-RELATED PROTEIN  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Disette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/773,368  
FILING DATE:  
CLASSIFICATION: 530  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy RJ  
REGISTRATION NUMBER: 36,749  
REFID: 415-845-4166  
TELEPHONE: 415-845-0555  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 266 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Geobank  
CLONE: 1030053  
; US-08-773-368-4

Sequence 10, Appli  
Sequence 6, Appli  
Sequence 7, Appli  
Sequence 8, Appli  
Sequence 5, Appli  
Sequence 10, Appli  
Sequence 11, Appli  
Sequence 4, Appli  
Sequence 29, Appli  
Sequence 28, Appli  
Sequence 30, Appli  
Sequence 31, Appli  
Sequence 32, Appli  
Sequence 33, Appli  
Sequence 34, Appli  
Sequence 35, Appli  
Sequence 36, Appli  
Sequence 37, Appli  
Sequence 38, Appli  
Sequence 39, Appli  
Sequence 40, Appli  
Sequence 41, Appli  
Sequence 42, Appli  
Sequence 43, Appli  
Sequence 44, Appli  
Sequence 45, Appli  
Sequence 46, Appli  
Sequence 47, Appli  
Sequence 48, Appli  
Sequence 49, Appli  
Sequence 50, Appli  
Sequence 51, Appli  
Sequence 52, Appli  
Sequence 53, Appli  
Sequence 54, Appli  
Sequence 55, Appli  
Sequence 56, Appli  
Sequence 57, Appli  
Sequence 58, Appli  
Sequence 59, Appli  
Sequence 60, Appli  
Sequence 61, Appli  
Sequence 62, Appli  
Sequence 63, Appli  
Sequence 64, Appli  
Sequence 65, Appli  
Sequence 66, Appli  
Sequence 67, Appli  
Sequence 68, Appli  
Sequence 69, Appli  
Sequence 70, Appli  
Sequence 71, Appli  
Sequence 72, Appli  
Sequence 73, Appli  
Sequence 74, Appli  
Sequence 75, Appli  
Sequence 76, Appli  
Sequence 77, Appli  
Sequence 78, Appli  
Sequence 79, Appli  
Sequence 80, Appli  
Sequence 81, Appli  
Sequence 82, Appli  
Sequence 83, Appli  
Sequence 84, Appli  
Sequence 85, Appli  
Sequence 86, Appli  
Sequence 87, Appli  
Sequence 88, Appli  
Sequence 89, Appli  
Sequence 90, Appli  
Sequence 91, Appli  
Sequence 92, Appli  
Sequence 93, Appli  
Sequence 94, Appli  
Sequence 95, Appli  
Sequence 96, Appli  
Sequence 97, Appli  
Sequence 98, Appli  
Sequence 99, Appli  
Sequence 100, Appli  
Sequence 101, Appli  
Sequence 102, Appli  
Sequence 103, Appli  
Sequence 104, Appli  
Sequence 105, Appli  
Sequence 106, Appli  
Sequence 107, Appli  
Sequence 108, Appli  
Sequence 109, Appli  
Sequence 110, Appli  
Sequence 111, Appli  
Sequence 112, Appli  
Sequence 113, Appli  
Sequence 114, Appli  
Sequence 115, Appli  
Sequence 116, Appli  
Sequence 117, Appli  
Sequence 118, Appli  
Sequence 119, Appli  
Sequence 120, Appli  
Sequence 121, Appli  
Sequence 122, Appli  
Sequence 123, Appli  
Sequence 124, Appli  
Sequence 125, Appli  
Sequence 126, Appli  
Sequence 127, Appli  
Sequence 128, Appli  
Sequence 129, Appli  
Sequence 130, Appli  
Sequence 131, Appli  
Sequence 132, Appli  
Sequence 133, Appli  
Sequence 134, Appli  
Sequence 135, Appli  
Sequence 136, Appli  
Sequence 137, Appli  
Sequence 138, Appli  
Sequence 139, Appli  
Sequence 140, Appli  
Sequence 141, Appli  
Sequence 142, Appli  
Sequence 143, Appli  
Sequence 144, Appli  
Sequence 145, Appli  
Sequence 146, Appli  
Sequence 147, Appli  
Sequence 148, Appli  
Sequence 149, Appli  
Sequence 150, Appli  
Sequence 151, Appli  
Sequence 152, Appli  
Sequence 153, Appli  
Sequence 154, Appli  
Sequence 155, Appli  
Sequence 156, Appli  
Sequence 157, Appli  
Sequence 158, Appli  
Sequence 159, Appli  
Sequence 160, Appli  
Sequence 161, Appli  
Sequence 162, Appli  
Sequence 163, Appli  
Sequence 164, Appli  
Sequence 165, Appli  
Sequence 166, Appli  
Sequence 167, Appli  
Sequence 168, Appli  
Sequence 169, Appli  
Sequence 170, Appli  
Sequence 171, Appli  
Sequence 172, Appli  
Sequence 173, Appli  
Sequence 174, Appli  
Sequence 175, Appli  
Sequence 176, Appli  
Sequence 177, Appli  
Sequence 178, Appli  
Sequence 179, Appli  
Sequence 180, Appli  
Sequence 181, Appli  
Sequence 182, Appli  
Sequence 183, Appli  
Sequence 184, Appli  
Sequence 185, Appli  
Sequence 186, Appli  
Sequence 187, Appli  
Sequence 188, Appli  
Sequence 189, Appli  
Sequence 190, Appli  
Sequence 191, Appli  
Sequence 192, Appli  
Sequence 193, Appli  
Sequence 194, Appli  
Sequence 195, Appli  
Sequence 196, Appli  
Sequence 197, Appli  
Sequence 198, Appli  
Sequence 199, Appli  
Sequence 200, Appli  
Sequence 201, Appli  
Sequence 202, Appli  
Sequence 203, Appli  
Sequence 204, Appli  
Sequence 205, Appli  
Sequence 206, Appli  
Sequence 207, Appli  
Sequence 208, Appli  
Sequence 209, Appli  
Sequence 210, Appli  
Sequence 211, Appli  
Sequence 212, Appli  
Sequence 213, Appli  
Sequence 214, Appli  
Sequence 215, Appli  
Sequence 216, Appli  
Sequence 217, Appli  
Sequence 218, Appli  
Sequence 219, Appli  
Sequence 220, Appli  
Sequence 221, Appli  
Sequence 222, Appli  
Sequence 223, Appli  
Sequence 224, Appli  
Sequence 225, Appli  
Sequence 226, Appli  
Sequence 227, Appli  
Sequence 228, Appli  
Sequence 229, Appli  
Sequence 230, Appli  
Sequence 231, Appli  
Sequence 232, Appli  
Sequence 233, Appli  
Sequence 234, Appli  
Sequence 235, Appli  
Sequence 236, Appli  
Sequence 237, Appli  
Sequence 238, Appli  
Sequence 239, Appli  
Sequence 240, Appli  
Sequence 241, Appli  
Sequence 242, Appli  
Sequence 243, Appli  
Sequence 244, Appli  
Sequence 245, Appli  
Sequence 246, Appli  
Sequence 247, Appli  
Sequence 248, Appli  
Sequence 249, Appli  
Sequence 250, Appli  
Sequence 251, Appli  
Sequence 252, Appli  
Sequence 253, Appli  
Sequence 254, Appli  
Sequence 255, Appli  
Sequence 256, Appli  
Sequence 257, Appli  
Sequence 258, Appli  
Sequence 259, Appli  
Sequence 260, Appli  
Sequence 261, Appli  
Sequence 262, Appli  
Sequence 263, Appli  
Sequence 264, Appli  
Sequence 265, Appli  
Sequence 266, Appli  
Sequence 267, Appli  
Sequence 268, Appli  
Sequence 269, Appli  
Sequence 270, Appli  
Sequence 271, Appli  
Sequence 272, Appli  
Sequence 273, Appli  
Sequence 274, Appli  
Sequence 275, Appli  
Sequence 276, Appli  
Sequence 277, Appli  
Sequence 278, Appli  
Sequence 279, Appli  
Sequence 280, Appli  
Sequence 281, Appli  
Sequence 282, Appli  
Sequence 283, Appli  
Sequence 284, Appli  
Sequence 285, Appli  
Sequence 286, Appli  
Sequence 287, Appli  
Sequence 288, Appli  
Sequence 289, Appli  
Sequence 290, Appli  
Sequence 291, Appli  
Sequence 292, Appli  
Sequence 293, Appli  
Sequence 294, Appli  
Sequence 295, Appli  
Sequence 296, Appli  
Sequence 297, Appli  
Sequence 298, Appli  
Sequence 299, Appli  
Sequence 300, Appli  
Sequence 301, Appli  
Sequence 302, Appli  
Sequence 303, Appli  
Sequence 304, Appli  
Sequence 305, Appli  
Sequence 306, Appli  
Sequence 307, Appli  
Sequence 308, Appli  
Sequence 309, Appli  
Sequence 310, Appli  
Sequence 311, Appli  
Sequence 312, Appli  
Sequence 313, Appli  
Sequence 314, Appli  
Sequence 315, Appli  
Sequence 316, Appli  
Sequence 317, Appli  
Sequence 318, Appli  
Sequence 319, Appli  
Sequence 320, Appli  
Sequence 321, Appli  
Sequence 322, Appli  
Sequence 323, Appli  
Sequence 324, Appli  
Sequence 325, Appli  
Sequence 326, Appli  
Sequence 327, Appli  
Sequence 328, Appli  
Sequence 329, Appli  
Sequence 330, Appli  
Sequence 331, Appli  
Sequence 332, Appli  
Sequence 333, Appli  
Sequence 334, Appli  
Sequence 335, Appli  
Sequence 336, Appli  
Sequence 337, Appli  
Sequence 338, Appli  
Sequence 339, Appli  
Sequence 340, Appli  
Sequence 341, Appli  
Sequence 342, Appli  
Sequence 343, Appli  
Sequence 344, Appli  
Sequence 345, Appli  
Sequence 346, Appli  
Sequence 347, Appli  
Sequence 348, Appli  
Sequence 349, Appli  
Sequence 350, Appli  
Sequence 351, Appli  
Sequence 352, Appli  
Sequence 353, Appli  
Sequence 354, Appli  
Sequence 355, Appli  
Sequence 356, Appli  
Sequence 357, Appli  
Sequence 358, Appli  
Sequence 359, Appli  
Sequence 360, Appli  
Sequence 361, Appli  
Sequence 362, Appli  
Sequence 363, Appli  
Sequence 364, Appli  
Sequence 365, Appli  
Sequence 366, Appli  
Sequence 367, Appli  
Sequence 368, Appli  
Sequence 369, Appli  
Sequence 370, Appli  
Sequence 371, Appli  
Sequence 372, Appli  
Sequence 373, Appli  
Sequence 374, Appli  
Sequence 375, Appli  
Sequence 376, Appli  
Sequence 377, Appli  
Sequence 378, Appli  
Sequence 379, Appli  
Sequence 380, Appli  
Sequence 381, Appli  
Sequence 382, Appli  
Sequence 383, Appli  
Sequence 384, Appli  
Sequence 385, Appli  
Sequence 386, Appli  
Sequence 387, Appli  
Sequence 388, Appli  
Sequence 389, Appli  
Sequence 390, Appli  
Sequence 391, Appli  
Sequence 392, Appli  
Sequence 393, Appli  
Sequence 394, Appli  
Sequence 395, Appli  
Sequence 396, Appli  
Sequence 397, Appli  
Sequence 398, Appli  
Sequence 399, Appli  
Sequence 400, Appli  
Sequence 401, Appli  
Sequence 402, Appli  
Sequence 403, Appli  
Sequence 404, Appli  
Sequence 405, Appli  
Sequence 406, Appli  
Sequence 407, Appli  
Sequence 408, Appli  
Sequence 409, Appli  
Sequence 410, Appli  
Sequence 411, Appli  
Sequence 412, Appli  
Sequence 413, Appli  
Sequence 414, Appli  
Sequence 415, Appli  
Sequence 416, Appli  
Sequence 417, Appli  
Sequence 418, Appli  
Sequence 419, Appli  
Sequence 420, Appli  
Sequence 421, Appli  
Sequence 422, Appli  
Sequence 423, Appli  
Sequence 424, Appli  
Sequence 425, Appli  
Sequence 426, Appli  
Sequence 427, Appli  
Sequence 428, Appli  
Sequence 429, Appli  
Sequence 430, Appli  
Sequence 431, Appli  
Sequence 432, Appli  
Sequence 433, Appli  
Sequence 434, Appli  
Sequence 435, Appli  
Sequence 436, Appli  
Sequence 437, Appli  
Sequence 438, Appli  
Sequence 439, Appli  
Sequence 440, Appli  
Sequence 441, Appli  
Sequence 442, Appli  
Sequence 443, Appli  
Sequence 444, Appli  
Sequence 445, Appli  
Sequence 446, Appli  
Sequence 447, Appli  
Sequence 448, Appli  
Sequence 449, Appli  
Sequence 450, Appli  
Sequence 451, Appli  
Sequence 452, Appli  
Sequence 453, Appli  
Sequence 454, Appli  
Sequence 455, Appli  
Sequence 456, Appli  
Sequence 457, Appli  
Sequence 458, Appli  
Sequence 459, Appli  
Sequence 460, Appli  
Sequence 461, Appli  
Sequence 462, Appli  
Sequence 463, Appli  
Sequence 464, Appli  
Sequence 465, Appli  
Sequence 466, Appli  
Sequence 467, Appli  
Sequence 468, Appli  
Sequence 469, Appli  
Sequence 470, Appli  
Sequence 471, Appli  
Sequence 472, Appli  
Sequence 473, Appli  
Sequence 474, Appli  
Sequence 475, Appli  
Sequence 476, Appli  
Sequence 477, Appli  
Sequence 478, Appli  
Sequence 479, Appli  
Sequence 480, Appli  
Sequence 481, Appli  
Sequence 482, Appli  
Sequence 483, Appli  
Sequence 484, Appli  
Sequence 485, Appli  
Sequence 486, Appli  
Sequence 487, Appli  
Sequence 488, Appli  
Sequence 489, Appli  
Sequence 490, Appli  
Sequence 491, Appli  
Sequence 492, Appli  
Sequence 493, Appli  
Sequence 494, Appli  
Sequence 495, Appli  
Sequence 496, Appli  
Sequence 497, Appli  
Sequence 498, Appli  
Sequence 499, Appli  
Sequence 500, Appli  
Sequence 501, Appli  
Sequence 502, Appli  
Sequence 503, Appli  
Sequence 504, Appli  
Sequence 505, Appli  
Sequence 506, Appli  
Sequence 507, Appli  
Sequence 508, Appli  
Sequence 509, Appli  
Sequence 510, Appli  
Sequence 511, Appli  
Sequence 512, Appli  
Sequence 513, Appli  
Sequence 514, Appli  
Sequence 515, Appli  
Sequence 516, Appli  
Sequence 517, Appli  
Sequence 518, Appli  
Sequence 519, Appli  
Sequence 520, Appli  
Sequence 521, Appli  
Sequence 522, Appli  
Sequence 523, Appli  
Sequence 524, Appli  
Sequence 525, Appli  
Sequence 526, Appli  
Sequence 527, Appli  
Sequence 528, Appli  
Sequence 529, Appli  
Sequence 530, Appli  
Sequence 531, Appli  
Sequence 532, Appli  
Sequence 533, Appli  
Sequence 534, Appli  
Sequence 535, Appli  
Sequence 536, Appli  
Sequence 537, Appli  
Sequence 538, Appli  
Sequence 539, Appli  
Sequence 540, Appli  
Sequence 541, Appli  
Sequence 542, Appli  
Sequence 543, Appli  
Sequence 544, Appli  
Sequence 545, Appli  
Sequence 546, Appli  
Sequence 547, Appli  
Sequence 548, Appli  
Sequence 549, Appli  
Sequence 550, Appli  
Sequence 551, Appli  
Sequence 552, Appli  
Sequence 553, Appli  
Sequence 554, Appli  
Sequence 555, Appli  
Sequence 556, Appli  
Sequence 557, Appli  
Sequence 558, Appli  
Sequence 559, Appli  
Sequence 560, Appli  
Sequence 561, Appli  
Sequence 562, Appli  
Sequence 563, Appli  
Sequence 564, Appli  
Sequence 565, Appli  
Sequence 566, Appli  
Sequence 567, Appli  
Sequence 568, Appli  
Sequence 569, Appli  
Sequence 570, Appli  
Sequence 571, Appli  
Sequence 572, Appli  
Sequence 573, Appli  
Sequence 574, Appli  
Sequence 575, Appli  
Sequence 576, Appli  
Sequence 577, Appli  
Sequence 578, Appli  
Sequence 579, Appli  
Sequence 580, Appli  
Sequence 581, Appli  
Sequence 582, Appli  
Sequence 583, Appli  
Sequence 584, Appli  
Sequence 585, Appli  
Sequence 586, Appli  
Sequence 587, Appli  
Sequence 588, Appli  
Sequence 589, Appli  
Sequence 590, Appli  
Sequence 591, Appli  
Sequence 592, Appli  
Sequence 593, Appli  
Sequence 594, Appli  
Sequence 595, Appli  
Sequence 596, Appli  
Sequence 597, Appli  
Sequence 598, Appli  
Sequence 599, Appli  
Sequence 600, Appli  
Sequence 601, Appli  
Sequence 602, Appli  
Sequence 603, Appli  
Sequence 604, Appli  
Sequence 605, Appli  
Sequence 606, Appli  
Sequence 607, Appli  
Sequence 608, Appli  
Sequence 609, Appli  
Sequence 610, Appli  
Sequence 611, Appli  
Sequence 612, Appli  
Sequence 613, Appli  
Sequence 614, Appli  
Sequence 615, Appli  
Sequence 616, Appli  
Sequence 617, Appli  
Sequence 618, Appli  
Sequence 619, Appli  
Sequence 620, Appli  
Sequence 621, Appli  
Sequence 622, Appli  
Sequence 623, Appli  
Sequence 624, Appli  
Sequence 625, Appli  
Sequence 626, Appli  
Sequence 627, Appli  
Sequence 628, Appli  
Sequence 629, Appli  
Sequence 630, Appli  
Sequence 631, Appli  
Sequence 632, Appli  
Sequence 633, Appli  
Sequence 634, Appli  
Sequence 635, Appli  
Sequence 636, Appli  
Sequence 637, Appli  
Sequence 638, Appli  
Sequence 639, Appli  
Sequence 640, Appli  
Sequence 641, Appli  
Sequence 642, Appli  
Sequence 643, Appli  
Sequence 644, Appli  
Sequence 645, Appli  
Sequence 646, Appli  
Sequence 647, Appli  
Sequence 648, Appli  
Sequence 649, Appli  
Sequence 650, Appli  
Sequence 651, Appli  
Sequence 652, Appli  
Sequence 653, Appli  
Sequence 654, Appli  
Sequence 655, Appli  
Sequence 656, Appli  
Sequence 657, Appli  
Sequence 658, Appli  
Sequence 659, Appli  
Sequence 660, Appli  
Sequence 661, Appli  
Sequence 662, Appli  
Sequence 663, Appli  
Sequence 664, Appli  
Sequence 665, Appli  
Sequence 666, Appli  
Sequence 667, Appli  
Sequence 668, Appli  
Sequence 669, Appli  
Sequence 670, Appli  
Sequence 671, Appli  
Sequence 672, Appli  
Sequence 673, Appli  
Sequence 674, Appli  
Sequence 675, Appli  
Sequence 676, Appli  
Sequence 677, Appli  
Sequence 678, Appli  
Sequence 679, Appli  
Sequence 680, Appli  
Sequence 681, Appli  
Sequence 682, Appli  
Sequence 683, Appli  
Sequence 684, Appli  
Sequence 685, Appli  
Sequence 686, Appli  
Sequence 687, Appli  
Sequence 688, Appli  
Sequence 689, Appli  
Sequence 690, Appli  
Sequence 691, Appli  
Sequence 692, Appli  
Sequence 693, Appli  
Sequence 694, Appli  
Sequence 695, Appli  
Sequence 696, Appli  
Sequence 697, Appli  
Sequence 698, Appli  
Sequence 699, Appli  
Sequence 700, Appli  
Sequence 701, Appli  
Sequence 702, Appli  
Sequence 703, Appli  
Sequence 704, Appli  
Sequence 705, Appli  
Sequence 706, Appli  
Sequence 707, Appli  
Sequence 708, Appli  
Sequence 709, Appli  
Sequence 710, Appli  
Sequence 711, Appli  
Sequence 712, Appli  
Sequence 713, Appli  
Sequence 714, Appli  
Sequence 715, Appli  
Sequence 716, Appli  
Sequence 717, Appli  
Sequence 718, Appli  
Sequence 719, Appli  
Sequence 720, Appli  
Sequence 721, Appli  
Sequence 722, Appli  
Sequence 723, Appli  
Sequence 724, Appli  
Sequence 725, Appli  
Sequence 726, Appli  
Sequence 727, Appli  
Sequence 728, Appli  
Sequence 729, Appli  
Sequence 730, Appli  
Sequence 731, Appli  
Sequence 732, Appli  
Sequence 733, Appli  
Sequence 734, Appli  
Sequence 735, Appli  
Sequence 736, Appli  
Sequence 737, Appli  
Sequence 738, Appli  
Sequence 739, Appli  
Sequence 740, Appli  
Sequence 741, Appli  
Sequence 742, Appli  
Sequence 743, Appli  
Sequence 744, Appli  
Sequence 745, Appli  
Sequence 746, Appli  
Sequence 747, Appli  
Sequence 748, Appli  
Sequence 749, Appli  
Sequence 750, Appli  
Sequence 751, Appli  
Sequence 752, Appli  
Sequence 753, Appli  
Sequence 754, Appli  
Sequence 755, Appli  
Sequence 756, Appli  
Sequence 757, Appli  
Sequence 758, Appli  
Sequence 759, Appli  
Sequence 760, Appli  
Sequence 761, Appli  
Sequence 762, Appli  
Sequence 763, Appli  
Sequence 764, Appli  
Sequence 765, Appli  
Sequence 766, Appli  
Sequence 767, Appli  
Sequence 768, Appli  
Sequence 769, Appli  
Sequence 770, Appli  
Sequence 771, Appli  
Sequence 772, Appli  
Sequence 773, Appli  
Sequence 774, Appli  
Sequence 775, Appli  
Sequence 776, Appli

Best Local Similarity 35.7%; Pred. No. 9.2e-25; Mismatches 75; Indels 40; Gaps 12; Matches 82; Conservative 33; Mismatches 75; Indels 40; Gaps 12;

Query Match 12.9%; Score 355; DB 3; Length 266; Best Local Similarity 35.7%; Pred. No. 9.2e-25; Mismatches 82; Conservative 33; Mismatches 75; Indels 40; Gaps 12; Matches 82; Conservative 33; Mismatches 75; Indels 40; Gaps 12;

QY 25 LPNVTLLEELSKYQHNEHSRVRRAIPRED-KEELMLHNLKLRGQVOPQASNMEMTWD 83  
Db 5 LATIWAQWASVFSVSYHT--ANTLPDIEFDFIKDCVRHNRKRSSEVKTAASDMLYMTWD 61  
Db 5 LATIWAQWASVFSVSYHT--ANTLPDIEFDFIKDCVRHNRKRSSEVKTAASDMLYMTWD 61  
QY 84 DELEKSAAAWASO�IWIHG---PTSL---LVSIGONLGAHMRYRSPGFHQS---W 131  
Db 62 PALAQIAKAWASNCQFSHNTRLKPPHKLHPNFTSILENI--W-TGSVPIFSVSSAINW 117  
QY 132 YDEVKDVTYPYSECNPWCPERCSPGPMCTHYOTIWATNTKIGCAVNTRKMTWGEWE 191  
Db 118 YDEIQDVF-----KTRICKKVCGHITOQVWADSYKVGCVAQFCPKVSGF-DALS 166  
QY 192 NAVYFVNYSPIGN--WIGEAPYKNGRPCSECPPSYGGSCRNLYCREE 238  
Db 118 YDEIQDVF-----KTRICKKVCGHITOQVWADSYKVGCVAQFCPKVSGF-DALS 166  
Db 167 NGAHFICNYPGCGNYPWT--PYKGATGTCACPN--DKCLDNLCVNRO 210

RESULT 2  
US-09-199-887-4  
; Sequence 4, Application US/0919987  
; Patent No. 6071874

GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Goli, Surya K.  
; TITLE OF INVENTION: NOVEL HUMAN PATHOGENESIS-RELATED  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Inoyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/199, 887  
; FILING DATE:  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/773, 368  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy RJ  
; REGISTRATION NUMBER: 36, 749  
; REFERENCE/DOCKET NUMBER: PF-0186 US

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 219 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 1030053  
; US-09-199-887-4

RESULT 3  
US-08-773-368-3  
; Sequence 3, Application US/08773368  
; Patent No. 5856130

GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Goli, Surya K.  
; TITLE OF INVENTION: NOVEL HUMAN PATHOGENESIS-RELATED  
; NUMBER OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Inoyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/773, 368  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy RJ  
; REGISTRATION NUMBER: 36, 749  
; REFERENCE/DOCKET NUMBER: PF-0186 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 219 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 847722  
; US-08-773-368-3

Query Match 12.8%; Score 352; DB 2; Length 219;  
; Best Local Similarity 37.0%; Pred. No. 1.3e-24;  
; Matches 81; Conservative 31; Mismatches 67; Indels 40; Gaps 12;

QY 35 ISKYOHNEHSRVRRAIPRED-KEELMLHNLKLRGQVOPQASNMEMTWD 93  
Db 5 VSNSHT--ANTLPDIEFDFIKDCVRHNRKRSSEVKTAASDMLYMTWDPALAQIAKAW 61

RESULT 4  
US-09-199-887-3  
; Sequence 3, Application US/09199887  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; TITLE OF INVENTION: NOVEL HUMAN PATHOGENESIS-RELATED  
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/199, 887  
; FILING DATE:  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/773, 368  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy RJ  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0186 US  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-845-0555  
; TELEFAX: 415-845-4166  
; TELLEX:  
; INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
; LENGTH: 219 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: Genbank  
; CLONE: 847722  
; US-09-199-887-3

Query Match 12.8%; Score 352; DB 3; Length 219;  
Best Local Similarity 37.0%; Pred. No. 1; 3e-24; Matches 81; Conservative 31; Mismatches 67; Indels 40; Gaps 12;

QY 35 LSIKYKHNHSHSRVRAATPREDKEELMLHNLKRGQVQPAASMEYMTWDELELSAAW 93  
Db 5 VSNVSHT--ANTLPDIENEDFKCVTRIHNRESEKPTASDMLYMTWDPLAQIAKAW 61  
QY 94 ASQC1WEHG----PTSL---LVSIGONLGAHNGRYRSPGFHWQS---WYDEVKDTYP 141

RESULT 5  
US-08-773-368-1  
; Sequence 1, Application US/08773368  
; Patent No. 5856130  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; TITLE OF INVENTION: NOVEL HUMAN PATHOGENESIS-RELATED  
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/773, 368  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy RJ  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0186 US  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELLEX:  
; INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
; LENGTH: 270 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: 1599164  
; US-08-773-368-1

Query Match 11.7%; Score 322; DB 2; Length 270;  
Best Local Similarity 35.6%; Pred. No. 1e-21; Matches 84; Conservative 27; Mismatches 105; Indels 20; Gaps 8;

QY 50 AIPREDKEELMLHNLKRGQVQPAASMEYMTWDELELSAAWASOC1WEHGCGTSLIVS 109  
Db 27 ALTDEEKRMLVNLHYRAQVSPASDMLMIRWDRLAFAKAYARQ-XRXGHNKGRR 84  
QY 110 IGGNLGAHNGRYRSPGFHWQSWDVKDTYPSECNCWPCECSPGMCTHYQIWT 169  
Db 85 RGENLFAITPEGMDPFLAMEWWHEHREHYNLS-RATCSP-----GOKMGHTQVWAK 136  
QY 170 TNKIGCAVNCRKMTWGEWENAVYFVCNYSPKGNIGEAPYKNGRCPOSECPSPGSC 229

Db 62 ASNCQFSHNTTRKPPHKLHPNFTSLGENI--W-TGSVPFLSVSSALTNWYDETDYNF- 116  
QY 142 YPSECNPWCPERCSPGMCTHYQIWTATNKIGCAVNCRKMTWGEWENAVYFVCNYS 201  
Db 117 -----KTRICKKVCGHYTQVWADSYKVGCNVQFCPKVSGF-DALSNGAHFTCNYG 166  
QY 202 PKGN--WIGEAPYKNGRPCSCCPPSYYGSCRNLYRE 237  
Db 167 PGGYPTW---PYKGATCSACPNN--DKCLDNLCVND 199

Db 137 TERIGCGSHFCEKQ--GVEETNELLVNYEPGNVKGKRPYQEGTPCSQCBSGY--HC 192  
 Qy 230 RNNICYREENTYPTPKETDEMNEVETAP1PEENIWLQPRM--RPPKKKISAVN 282  
 Db 193 KNSLC--EPIGSPEADAQDPLYVTEAPSFRATEA-SDSRKMGAEGPDKPSVSGLN 245

RESULT 6

US-09-199-887-1

Sequence 1, Application US/09199887

Patent No. 6071874

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Goli, Surya K.

TITLE OF INVENTION: NOVEL HUMAN PATHOGENESIS-RELATED

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Indyc Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ FOR Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/199,887

FILING DATE:

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/773,368

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy RJ

REFERENCE/DOCKET NUMBER: 36,749

TELECOMMUNICATION INFORMATION:

TELEPHONE: (310) 863-0223

TELEFAX: (303) 863-9700

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2618-30-PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (310) 863-9700

TELEFAX: (303) 863-9700

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2618-30-PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (310) 863-9700

TELEFAX: (303) 863-9700

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2618-30-PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (310) 863-9700

TELEFAX: (303) 863-9700

Query Match 11.7%; Score 322; DB 3; Length 270;  
 Best Local Similarity 35.6%; Pred. No. 1e-21;  
 Matches 84; Conservative 27; Mismatches 105; Indels 20; Gaps 8;

Qy 50 AIPRDEKEELMLHNLKLRGQVQPOSANMEYMTWDEBLEKSAAWASQCIWEHGPSSLVS 109  
 Db 27 ALTDEKEELMLHNLKLRGQVQPOSANMEYMTWDEBLEKSAAWASQCIWEHGPSSLVS 109  
 Qy 110 IGGNIGGAHWGRYRSFGHVSQWDYKVDPYVPSSECNPWCPCRGSPCMTHYQIVWAT 169  
 Db 85 RGENELAITDEGMDYPLAMEWHRHREHYNLS-AATCSP-----GQCGHIVQWVAK 136  
 Qy 170 TNKIGCAVNICKMWTGEMWENAVYEVNISPKGNWIEAPYNGRSPSECPSYGGSC 229  
 Db 137 TERICGSHFCEKLQ--GVEETNELLVNYEPGNVKGKRPQEGTCSCPSGY--HC 192  
 Qy 230 RNNICYREENTYPTPKETDEMNEVETAP1PEENIWLQPRM--RPPKKKISAVN 282

RESULT 8

PCT-US96-07709-25

Sequence 25, Application PC/TUS9607709

GENERAL INFORMATION:

APPLICANT: Tripp, Cynthia A.

APPLICANT: Wisniewski, Nancy

TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN

TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS

Db 193 KNSLC--EPIGSPEADAQDPLYVTEAPSFRATEA-SDSRKMGAEGPDKPSVSGLN 245  
 RESULT 7  
 PCT-US96-07709-30

Sequence 30, Application PC/TUS9607709

GENERAL INFORMATION:

APPLICANT: Tripp, Cynthia A.

APPLICANT: Wisniewski, Nancy

TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN

TITLE OF INVENTION: ANTIGEN 5-LINE GENES AND PROTEINS

Query Match 9.9%; Score 273; DB 5; Length 220;  
 Best Local Similarity 29.5%; Pred. No. 2.5e-17;  
 Matches 70; Conservative 36; Mismatches 85; Indels 46; Gaps 10;

Qy 16 IVCSSQGYLIF--NVTLEPLSLSKYQHNNESSHRSYRRATPREDEKEELMLHNLKLR--GQV 70  
 Db 10 IWWAVTYGNGFGKLTALER-----KKIVGQNKYRSIDLINGLKLKNRGT 55  
 Qy 71 QPOASNMENMTWDEBLEKSAAWASQCIWEHGPSSLVSIGGLNQAHM-----GRYRSP 124  
 Db 56 MPGRKMLRLWDCKLESSAQWRANOCITFGHSRQRQREGVGENVAYWSSVSEGLKTA 115  
 Qy 125 GFHV-QSWDEVKDYTYPSPCNCWPWCPCRGSPCMTHYQIVWATNKKICAVNT-C-- 180  
 Db 116 GTDACKSWNSLPLRYENNPNINNTW--KVAQGQVLHFQHAWKGTYKIGGVATODC 172  
 Qy 181 -RKTVWGVWENAVYEVNISPKGNWIEAPYNGRSPSECPSYGGSC--RNNIC 234  
 Db 173 GRTLIV-----ICHYSPGNNMVGEVYIORGNPCKVVKDCYTKKCLSKSLGC 218

RESULT 8

PCT-US96-07709-25

Sequence 25, Application PC/TUS9607709

GENERAL INFORMATION:

APPLICANT: Tripp, Cynthia A.

APPLICANT: Wisniewski, Nancy

TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN

TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS

NUMBER OF SEQUENCES: 43  
 CORRESPONDENCE ADDRESS: Sheridan Ross & McIntosh  
 ADDRESSEE: Sheridan Ross & McIntosh  
 STREET: 1700 Lincoln St., Suite 3500  
 CITY: Denver  
 STATE: Colorado  
 COUNTRY: U.S.A.  
 ZIP: 80203  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: IBM PC compatible  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US96/07709  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US96/07709  
 FILING DATE: 23-MAY-1996  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Connell, Gary J.  
 REGISTRATION NUMBER: 32,020  
 REFERENCE/DOCKET NUMBER: 2618-30-PCT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303) 863-9700  
 TELEFAX: (303) 863-0223  
 INFORMATION FOR SEQ ID NO: 25:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 248 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 PCT-US96-07709-25

NAME: Connell, Gary J.  
 REGISTRATION NUMBER: 32,020  
 REFERENCE/DOCKET NUMBER: 2618-30-PCT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303) 863-9700  
 TELEFAX: (303) 863-0223  
 INFORMATION FOR SEQ ID NO: 25:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 248 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 PCT-US96-07709-33

Query Match 9.9%; Score 273; DB 5; Length 248;  
 Best Local Similarity 29.5%; Pred. No. 2.9e-17; Indels 46; Gaps 10;  
 Matches 70; Conservative 36; Mismatches 85; Length 248;

QY 16 LVGGSGQGTLP--NVTLEELLSKQHNEHSRVRRAIPREDEKEILMLHNKL--GQV 70  
 Db 38 IIVAVTGNCPGKLTALER-----KKIVGONKVKRSDLINGKLNRRNGT 83

QY 71 QPQASNMETWDDLELSAAMAWASOClWEGPTSLVLSVGNQGAH-----GRYRSP 124  
 Db 84 MPRGKNMELRWDCKLLESSAQRMNQACIFGSPROOREGVGENVIAWSSVSUGLKKA 143

QY 125 CFHV-QSWDEVKRYTYPSECPKNPWERCPSGPMCTHYQIWATNTKIGCAVNT-C-- 180  
 Db 144 GTDAGKSWMSELPKLYENRPSNMW--KVAGQGVHFTOMAWGTYKICGGVATQCDG 200

QY 181 -RKMTWGEWENAVYFVCNYSPKGNWIGEAPKNGRPCSCPPSYGGSC--RNNIC 234  
 Db 201 GHNLIV-----ICHYSPGGNMVGEVIIQRGNPCKVDKDCYTKKCLSKGLC 246

RESULT 9  
 PCT-US96-07709-33  
 Sequence 33 Application PC/PCT-US9607709  
 GENERAL INFORMATION:  
 APPLICANT: Tripp, Cynthia A.  
 APPLICANT: Wisniewski, Nancy  
 TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN  
 NUMBER OF SEQUENCES: 43  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sheridan Ross & McIntosh  
 STREET: 1700 Lincoln St., Suite 3500  
 CITY: Denver  
 STATE: Colorado  
 COUNTRY: U.S.A.  
 ZIP: 80203  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/419,414  
 FILING DATE: 08-AUG-1994  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pabst, Patrea L.  
 REGISTRATION NUMBER: 31,284

OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US96/07709  
 FILING DATE: 23-MAY-1996  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Connell, Gary J.  
 REGISTRATION NUMBER: 32,020  
 REFERENCE/DOCKET NUMBER: 2618-30-PCT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303) 863-9700  
 TELEFAX: (303) 863-0223  
 INFORMATION FOR SEQ ID NO: 33:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 204 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 PCT-US96-07709-33

Query Match 9.7%; Score 269; DB 5; Length 204;  
 Best Local Similarity 31.0%; Pred. No. 5.2e-17; Mismatches 76; Indels 30; Gaps 8;  
 Matches 63; Conservative 34; Mismatches 76;

QY 48 RFAIRPREDKELMLHNKL--GQVOPQASNMETWDDLELSAAMAWASOClWEGPT 104  
 Db 14 KRVGONKVKRSDLINGKLNRRNGTMYMRGKNMELRWDCKLLESSAQRMNQACIFGSPR 73

QY 105 SLIVSICQNLGAH-----GRYRSPFHV-QSMWDEVKDYYPPYSECNPWCPBCSGP 157  
 Db 74 QDREGVENVVAYWSSVSVEGLKTPAGTDAGKSWMSELPKLYENRPSNMW-----KVAGQ 130

QY 158 MCTHYQIWATNTKIGCAVNT-C--RKNMVWGEWENAVYFVCNYSPKGNWIGEAPK 213  
 Db 131 GVLFPTOMAWGTYKICGGVATQCDGRTLV-----ICHYSPGGNMVGEVIIQ 179

QY 214 NGRCOSECPSPSYGGSC--RNNIC 234  
 Db 180 RGNPCKVDKDCYTKKCLSKGLC 202

RESULT 10  
 US/08/419,414-9  
 Sequence 9, Application US/08419414  
 ; PATENT NO. 573787  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hawton, John M.  
 ; APPLICANT: Holtez, Peter J.  
 ; APPLICANT: Jones, Brian F.  
 ; TITLE OF INVENTION: Hookworm Vaccine  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Patrea L. Pabst  
 ; STREET: 2800 One Atlantic Center  
 ; STREET: 1201 West Peachtree Street  
 ; CITY: Atlanta  
 ; STATE: Georgia  
 ; COUNTRY: USA  
 ; ZIP: 30309-3450  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/419,414  
 FILING DATE: 08-AUG-1994  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pabst, Patrea L.  
 REGISTRATION NUMBER: 31,284



SEQUENCE CHARACTERISTICS:  
 LENGTH: 204 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLogy: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE:  
 ORIGINAL SOURCE:  
 ORGANISM: *Vespa vulgaris*  
 US-09-130-287-2

RESULT 13  
 US-07-930-686-12  
 Sequence 12, Application US/07930686  
 ; Patent No. 5525508  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sharp, Phillip J  
 ; APPLICANT: Wagland, Barry M  
 ; APPLICANT: Cobon, Gary S  
 ; TITLE OF INVENTION: Nematode Vaccine  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley and Lardner  
 ; STREET: suite 500, 1800 Diagonal Road  
 ; CITY: Alexandria  
 ; STATE: Virginia  
 ; COUNTRY: United States of America  
 ; ZIP: 22313-0399  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/930,686  
 ; FILING DATE: 1992/06/06  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: AU PK4486  
 ; FILING DATE: 06-FEB-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/AU92/00040  
 ; FILING DATE: 06-FEB-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bent, Stephen A  
 ; REGISTRATION NUMBER: 29,768  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 836-9300  
 ; TELEFAX: (703) 683-4109

Query Match 9.3%; Score 256; DB 3; Length 204;  
 Best Local Similarity 30.5%; Pred. No. 8.1e-16;  
 Matches 71; Conservative 33; Mismatches 73; Indels 56; Gaps 12;

QY 1 MSCVLLGGVPLGLLFLVCGSSQGILPNTVLEELSKYQHNEHSRSVRRAIPREDKEIL 60  
 Db 6 MLHKLKGGV-----HTAC-KYGSKPNCG--NKVVVS-----GLTKOEKODIL 45

QY 61 MLHKLKGVO-----QPOASNMEYMTWDELEKSAAWASOCITWHRGPISLLV 108  
 Db 46 KEHNDFRQKARGLETRGNQGPQPQPKNMKNLWNLDAVVAQWANOCQHQDTCRDA 105

QY 109 --SIGNOL--GAHNGRYRSPGFHYOSWYIDEVKDYTPYRSECNPWCPCERG--MCT 160  
 Db 106 KYQVQNVNVALTGSTAAKYDDPVKLYKMWEDEVKDY-----NP--KKKFGNDFLKTG 155

QY 161 HYIQIYWATINKIGCAVNTRKMTWGEWENAVVYVCNYSPKGNNWGEAPYK 213  
 Db 156 HYIQIYWATINKIGCAVNTRKMTWGEWENAVVYVCNYSPKGNNWGEAPYK 202

RESULT 14  
 US-08-460-998-12  
 Sequence 12, Application US/08460998  
 ; Patent No. 5942413  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sharp, Phillip J  
 ; APPLICANT: Wagland, Barry M  
 ; APPLICANT: Cobon, Gary S  
 ; TITLE OF INVENTION: Nematode Vaccine  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley and Lardner  
 ; STREET: suite 500, 3000 K Street, NW  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: United States of America  
 ; ZIP: 20007-5109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/460,998  
 ; FILING DATE: 05-JUN-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/930,686  
 ; FILING DATE: 06-OCT-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: AU PK4486  
 ; FILING DATE: 06-FEB-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/AU92/00040  
 ; FILING DATE: 06-FEB-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bent, Stephen A  
 ; REGISTRATION NUMBER: 29,768  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 836-9300  
 ; TELEFAX: (703) 683-4109

Query Match 8.9%; Score 244.5; DB 1; Length 440;  
 Best Local Similarity 29.7%; Pred. No. 2.8e-14;  
 Matches 73; Conservative 32; Mismatches 86; Indels 55; Gaps 10;

QY 39 QHNESRSRVRRAIPREDKEEILMLHKLKGVOQPOASNMEYMTWDELEKSAAWASOCIT 98  
 Db 32 KHNETRSIIAKG-----QAKNKL-GSFPAKARMLKVGYDCEVEANTRAYAKECK 80

QY 99 WEHGTSILSLSIGQNL---GAHWGRRSRPGPHYSQTYDEVKDYTYPSECNPWCFCERC 154  
 Db 81 FEHDPSQEQRNYWQNLWMGGTNTSYKTESAKLSVQAWYELKMFGVPDENILTMVFDRG 140

QY 155 SGPMCIHYIQIYWATINKIGCAVNTRKMTWGEWENAVVYVCNYSPKGNNWGEAPYK 214  
 Db 141 VG---HYIQIYWATINKIGCAVNTRKMTWGEWENAVVYVCNYSPKGNNWGEAPYK 186

QY 215 GRPCS---ECPPSYGGSCRNNLYCREETY----TPKPEDEMNVEETAP---- 256  
 Db 187 GDPCTIDEDC----QCTGCTCSKDEALCIPGYTIVMPPTIE-RPTTPKIHPGM 238

QY 257 IPEENH 262  
 Db 239 CPENN 244

TELEPHONE: (202) 672-5300  
 TELEFAX: (202) 672-5399  
 IN INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 440 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 ;  
 ; MOLECULE TYPE: protein  
 ; US-0-460-998-12

Query Match 8.9%; Score 244.5; DB 2; Length 440;  
 Best Local Similarity 29.7%; Pred. No. 2.8e-14; Mismatches 86; Indels 55; Gaps 10;  
 Matches 73; Conservative 32; MisMatches 86; ;  
 Qy 39 OHNEHSRVRRAIPREDEKEEILMLNKKLRQVQPPASMENTWDDLELEKSAAWASQCI 98  
 Db 32 KHNEYRSITIAGK-----OAKNL-GGPAPKARMLKVGYDCVEANVAAKECK 80  
 Qy 99 WEHGGPSLVLVSTGQNL---GAHWGRYRSPGPHVQSYDEVKDVTYPYPSECNPWC PERC 154  
 Db 81 FEHDPPBQRQNYWGQNLWMLGGINYSKTESAKLUSQAVIWEKLKMFYGPDENILTM EFDRG 140  
 Qy 155 SQPMCTHYTQIWATNKIGCAVNTCRKMTWGEWENNAVIFVCNTSPKGWIGEAPKN 214  
 Db 141 VG---HYTOPAWQSSDKIGCAVECPMTL-----VACEYNPAGNRINHYDI 186  
 Qy 215 GRPCS---ECPSYGGSCRNNLYCREF-----TPKPETDEMNEVETAP----- 256  
 Db 187 GDPCTADBC-----OCTGCTCSKDEALCIPPGYTYMPPT-----KPTTPKIVHPGM 238  
 Qy 257 IPEENH 262  
 Db 239 CPENN 244  
 ;

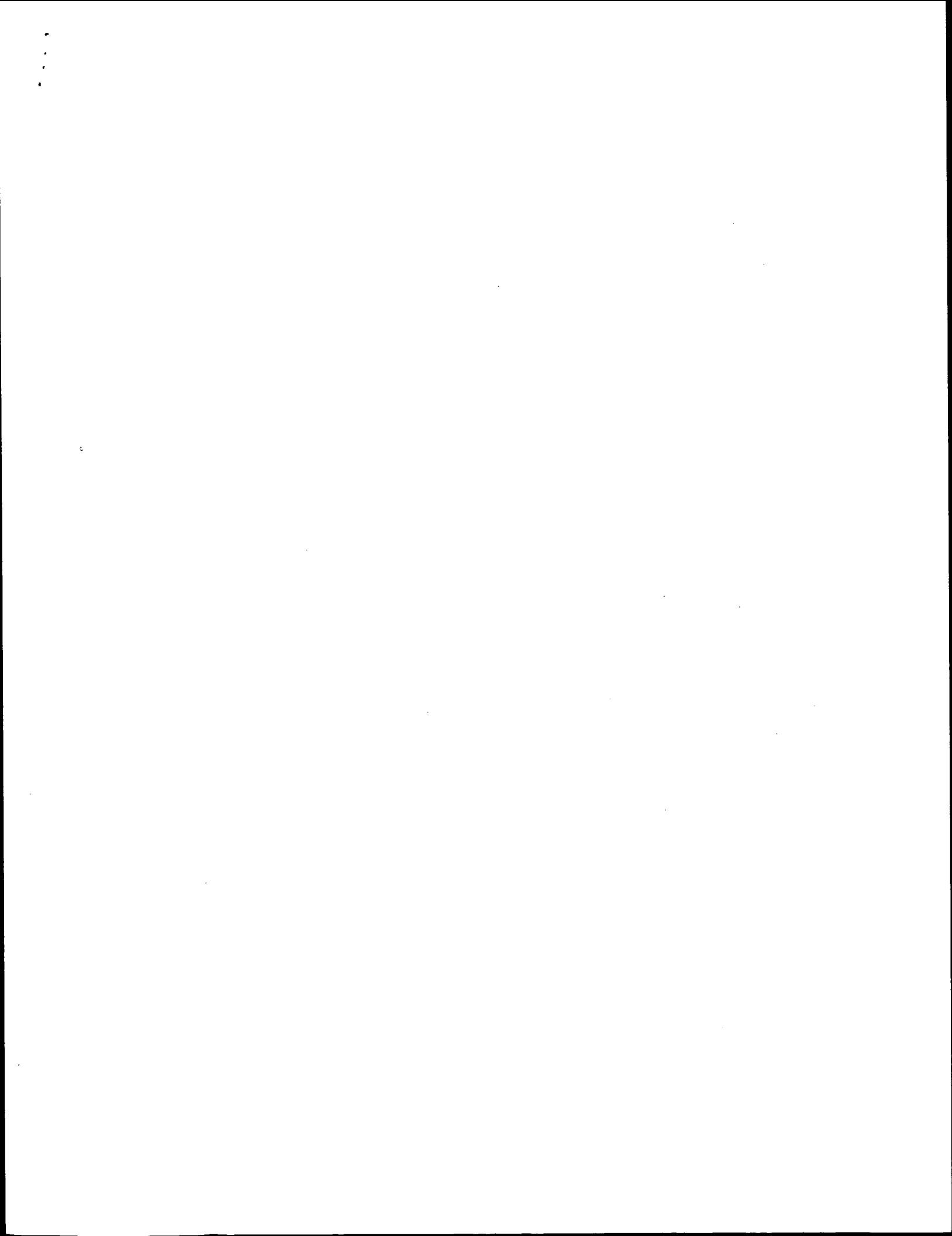
RESULT 15  
 US-08-614-935-1  
 ; Sequence 1, Application US/08614935  
 ; Patent No. 5804201  
 ; GENERAL INFORMATION:  
 ; APPLICANT: King, Te P.  
 ; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES OF VESPID  
 ; NUMBER OF SEQUENCES: 81  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: David A. Jackson, Esq.  
 ; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
 ; STREET: Floor  
 ; CITY: Hackensack  
 ; STATE: New Jersey  
 ; COUNTRY: USA  
 ; ZIP: 07601

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/614,935  
 FILING DATE: 11-MAR-1996  
 CLASSIFICATION: 436  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jackson Esq., David A.  
 REGISTRATION NUMBER: 26,742  
 REFERENCE/DOCKET NUMBER: 600-1-156  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201-487-5800  
 TELEFAX: 201-343-1684  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 204 amino acids  
 ; US-08-614-935-1

Query Match 8.7%; Score 240; DB 1; Length 204;  
 Best Local Similarity 29.6%; Pred. No. 2.4e-14; Mismatches 85; Indels 50; Gaps 9;  
 Matches 68; Conservative 27; MisMatches 85; ;  
 Qy 1 MSCVLLGGVIPPLGLFLYCGSGQYLPNVTLELLSKVQHNEHSRVRRAIPREDKEIL 60  
 Db 6 IKCLKGYY-----HTAC-KYSLKPGC-----GNKKVVSYGLPKOEDIL 45  
 Qy 61 MLINKLRCQV-----QPOQASMEYMWDDLELEKSAAWASQCIWEHGGPSLVL 108  
 Db 46 KEHNDFRKIAKGLETTRGNPGPQPOPKANNKLNWSDELAYIAQWANQOCQHGHTCDVA 105  
 Qy 109 -SIGONI---GAHWGRYRSPGPHVQSYDEVKDVTYPYPSECNPWC PERCSPGPMCHYT 163  
 Db 106 KYQVQNVNALTGSTAAVNDPYKLVKWNEDVKDINPKKSFENNFL-----KIGHYT 158  
 Qy 164 QTWATTNKIGCAVNTCRKMTWGEWENNAVIFVCNTSPKGWIGEAPKN 213  
 Db 159 QMWANTKEVGCG----SIKYIOENWHKH-YLVCNYGPSGNFONEELYQ 202  
 ;

Search completed: May 25, 2002, 14:28:50  
 Job time: 3316 sec







C;Accession: JCC4131  
 R;Murphy, E.V.; Zhang, Y.; Zhu, W.; Biggs, J.  
 Gene 159, 131-135, 1995  
 A;Title: The human glioma pathogenesis-related protein is structurally related to plant  
 A;Reference number: 91100927; PIDN:AAA82731.1; PID:9847722  
 A;Cross-references: GDB:683195  
 C;Superfamily: Yellowjacket venom allergen antigen 5  
 C;Keywords: brain

Query Match  
 Best Local Similarity 12.8%; Score 352; DB 2; length 219;

Matches 81; Conservative 31; Mismatches 67;

Indels 40; Gaps 12;

Db 5 LIVLFLLA-----VLP-----PSSLQDTDEWDRDLENLSTKLSV-----QEE-----GAHNGRY 121

Qy 35 LSKYQHNEHSRSRRVRAATPRED-KEEILMLNKLGVOPOASQMEWTMWDPELEKSAAW 93

Db 182 KMTVNGEVENAVAFVNVSPPKGNNWGE-----APYKNGRCSCSPSYGGSCRNCYRRT 239

Qy 94 ASOQIWEKG-----PRSL----LVSIGQIGAHMGRAWGRYRSGFHQS---WDEVKQTYP 141

Db 161 DQL-----KTVQHICPGGNVGRILSITEGCPCDCSCP----NCEDGLCNSCE 209

Qy 142 YFSECNPWMCERCSGPMTHIQIWIWATNKIGCAVNTCKTVAKEWVNAWYFVCWMS 201

Db 210 Y-----EDNY-----SNGDLKURKMSCDIDLKRG 235

Qy 117 -----KTRICKKVCGHQTQVWASVYKQVFCPVSG-----DALSNGAHCIONYG 166

Db 299 KGS 301

Qy 167 PGSNYPWV---PYKRGATCSCACPNN--DKLDLNLVND 199

RESULT 3

A24609 acidic epididymal glycoprotein

C;Species: sperm-coating glycoprotein - rat

C;Date: 30-Jun-1988 #sequence\_revision

R;Charest, N.J.; Joseph, D.R.; Wilson, E.M.; French, F.S.

A;Title: Molecular cloning of complementary deoxyribonucleic acid for an androgen-dependent cloning of complementary deoxyribonucleic acid for an androgen-regula

A;Accession: A40918; MUID:8909913

A;Status: preliminary

A;Residues: 1-246 <CHA>

R;Brooks, D.; Means, A.R.; Wright, E.J.; Singh, S.P.; Tiver, K.K.

A;Reference number: A24609; MUID:87053995

A;Molecule type: mRNA

A;Cross-references: GB:A0918

Bur. J. Biochem. 161, 13-18, 1986

A;Title: Molecular cloning of the DNA for androgen-dependent sperm-coating glycoprotein

A;Accession: A24609; MUID:87053995

A;Residues: 1-246 <CHA>

C;Superfamily: cysteine-rich secretory protein 1 (CRISP-1; DE/AEG) and the

C;Keywords: glycoprotein; sperm protein

F;1-19/Domain: signal sequence #status predicted <SIG>

Query Match  
 Best Local Similarity 10.9%; Score 301; DB 2; Length 246;  
 Matches 81; Conservative 51; Mismatches 90; Indels 81; Gaps 15;

OY 11 LGULFLVCGSSQGYLLPNTLLEELKSYQHNEHSRSRRVRAATPRED-KEEILMLNKLGVOPOASQMEWTMWDPELEKSAAW 93

Db 5 LVLFLLA-----VLP-----PSSLQDTDEWDRDLENLSTKLSV-----QEE-----GAHNGRY 121

Qy 71 OPOQASQMEWTMWDPELEKSAAWQSQCWHEQPLSVS---IGQIG-----AHNGRY 182

Db 54 SPSGSLKQWENNDQAVQWADQCTTSHPSPRLRTLNLRGENLMESYLAWSWSS-- 111

Qy 123 SPGFHFOSSWDEVKDVTPTVSECPWMCERCSGPMTHIQIWIWATNKIGCAVNTCKTVAKEWVNAWYFVCWMS 201

Db 112 -----AIGQWNEYKDLITY-----DVGPKQPDPSVW-----HTOVWNSFQVAGCVAECPK 159

Qy 183 MTWGEWENAVAFVNVSPPKGNNWGE-----APYKNGRCSCSPSYGGSCRNCYRRT 240

Db 160 NPL-----RYYVCHCPVGNIOGRLYTPYTAGEPASCPCPDHCDCDGICLTNSCGHEDKY 212

RESULT 5

S6891 neutrophil granules matrix glycoprotein SGP28 precursor - human

C;Species: Homo sapiens (man)  
C;Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Nov-2000  
C;Accession: S68691; S7413; S68683  
C;Status: Preliminary  
R;Kjeldsen, L.; Cowland, J.B.; Johnsen, A.H.; Borregaard, N.  
FBS Lett. 380, 246-250, 1996  
A;Title: SG28, a novel matrix glycoprotein in specific granules of human neutrophils w/  
A;Reference number: S68691; MUID:96186934  
A;Accession: S68691; mRNA  
A;Residues: 1-245 <KJL>  
A;Cross-references: EMBL:X94323; NID:91213612; PIDN:CAA63984.1; PID:91213613  
A;Accession: S74313  
A;Molecule type: protein  
A;Residues: 33-83; 96-143; 165-217; 221-226 <KJL>  
R;Kraetzschar, J.; Haendler, B.; Eberspaecher, U.; Roosterman, D.; Donner, P.; Schleunig, J.; Biochem. 236, 827-836, 1996  
A;Title: The human cysteine-rich secretory protein (CRISP) family. Primary structure and  
A;Reference number: S68681; MUID:96270732  
A;Accession: S68683  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Gene: SGP28  
A;Residues: 1-105 'S' 107-245 <KRN>  
A;Cross-references: EMBL:X92420; NID:91262818; PIDN:CAA64527.1; PID:91262819  
C;Genetics:  
C;Superfamily: cysteine-rich secretory protein 1  
C;Domain: signal sequence #status predicted <SIG>  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-245/Product: neutrophil granules matrix glycoprotein SGP28 #status predicted <MAT>  
Query Match 10.7%; Score 295.5; DB 2; Length 245;  
Best Local Similarity 31.0%; Pred. No. 1.9e-15;  
Matches 75; Conservative 29; Mismatches 91; Indels 47; Gaps 10;  
Query 13 LIFTVCGSGQYLLPNTVLLERLLSKYQHNEHSRSVRRAI--PREDKEEILMLHNLKRQ 69  
Db 7 LLFLVAG-----LLPSPEANEDKDPRTFTLITQTQVORELVNKHLRR 52  
Query 70 VQPQASNMETWDDLELSKAAAASOCIWEH-GPTSLVSI-GONL----GAHWGRY 121  
Db 53 VSPPARNLKMKWENKEAKAAQKWAQNCNYRHSNPKDRMTSLKGGENLYMSASSWSQ- 111  
Query 122 RSGFPFHQSWSIDEVKOTYPPSECNPWCPCRGPMCTHTQIWTATNNIGCAVNTCR 181  
Db 112 -----ATOSWDEYNDPFD---GVGPKTPNAVG---HYIQVWVWISSIVLGCGNAYCP 158  
Query 182 KMWVWGEWENAWAVFWCNYSPKGNWIG--APYKNGRPCSCCPPSPGGSERNLNLCYREE 239  
Db 159 NQKVL-----KYYVCQYPAGNWANRLYVPEEQGAPCASCPCDNCDDGLCTNGCRYEDL 212  
Query 240 YT 241  
Db 213 YS 214

## RESULT

6

B33329 cysteine-rich secretory protein 2 type I precursor - human  
N;Alternate names: testis-specific protein  
C;Species: Homo sapiens (man)  
C;Accession: B33329; S68682  
C;Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 20-Jun-2000  
C;Accession: B33329; S68682  
R;Kaschara, M.; Gutknecht, J.; Brew, K.; Spurr, N.; Goodfellow, P.N.  
Genomics 5, 521-534, 1989  
A;Title: Cloning and mapping of a testis-specific gene with sequence similarity to a spe  
A;Reference number: A33329; MUID:90129048

A;Accession: B33329  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Cross-references: GB:M25532; NID:9339882; PIDN:AAA61220.1; PID:9339883  
A;Cross-references: EMBL:X94323; NID:91213612; PIDN:CAA63984.1; PID:91213613  
R;Kraetzschar, J.; Haendler, B.; Eberspaecher, U.; Roosterman, D.; Donner, P.; Schleunig, J.; Biochem. 236, 827-836, 1996

A;Title: The human cysteine-rich secretory protein (CRISP) family. Primary structure  
A;Reference number: S68681; MUID:96270732  
A;Accession: S68682  
A;Status: Preliminary  
A;Molecule type: mRNA  
A;Residues: 1-243 <KRA>  
A;Cross-references: EMBL:X95239; NID:91262816; PIDN:CAA64526.1; PID:91262817  
C;Genetics:  
A;Gene: CDP:TPX1  
A;Cross-references: GDB:120760; OMIM:187430  
A;Map position: 6p21-6qter  
A;Superfamily: cysteine-rich secretory protein 1  
F;1-25/Domain: signal sequence #status predicted <SIG>  
F;21-243/Product: cysteine-rich secretory protein 2 type I #status predicted <MAT>  
Query Match 10.2%; Score 281; DB 2; Length 243;  
Best Local Similarity 26.3%; Pred. No. 2.5e-14;  
Matches 83; Conservative 33; Mismatches 83; Indels 116; Gaps 14;  
Query 24 LIPNVTLLEELLSKYQHNEHSRSVRRAI-----ELMLHNLKRQ 68  
Db 3 LIPVLELTVLP-----SLPAGKDPAFTALTTQLQVOREBVINKHLRK 49  
Query 69 QYPOQAASMEYMWDDLELSKAAAASOCIWEH-PTSLVNS-IGQNLGAHWGRYSP- 124  
Db 50 AVSPPPASMLKAEWSRSRVTTAQWRANKTLOHSDPFDKRTSTRGENL-----YMSSD 103  
Db 125 ---GFHIVOSWTDDEVKOTYPPSECNPWCPCRGPMCTHTQIWTATNNIGCAVNTC 180  
Query 104 PTSSWAIQSWSIDEVKOTYPPSECNPWCPCRGPMCTHTQIWTATNNIGCAVNTC 155  
Db 181 RKMIVWGEWENAWAVFWCNYSPKGNWIG--EAPYKNGRPCSCCPPSPGGSERNLNLCYREE 238  
Db 156 PNQD-----SLKVVYVQCICPAGNNMRKNTPYQGQTPCAGCP---DDDKGIC--- 201  
Query 239 TYTKPTEDEMEVETAPIPEEHVWQPRVMRPTKRTSAVNMYTQVVRDT----- 292  
Db 202 -----TNSCQYQDLISNCDSLKLNTAG 222  
Query 293 ---KMKDRKGSTC 303  
Db 223 CEHLLIKEKCK-ATC 236

## RESULT

7

T16415 hypothetical protein F48E8.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C;Accession: T16415

submitted to the EMBL Data Library, March 1995

R;Kirrten, J.  
submitted to the EMBL Data Library, March 1995

A;Description: The sequence of C. elegans cosmid F48E8.

A;Reference number: S59413

A;Accession: T16415  
A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-312 <KIR>

A;Cross-references: EMBL:U23514; NID:9746484; PID:9746485; PIDN:AAC46538.1; CESP:F48E

A;Experimental source: strain Bristol N2

C;Genetics:  
A;Gene: CESP:F48E8.1

A;Introns: 28/1; 94/3; 153/3; 174/3; 226/3

Query Match 9.7%; Score 269; DB 2; Length 312;  
Best Local Similarity 28.4%; Pred. No. 2.8e-13;  
Matches 83; Conservative 36; Mismatches 93; Indels 80; Gaps 17;

Query 21 QGYLLPVNLLEELLSKYQHNE-----SHSRVRAI-----ELMLHNLKRQ 70

Db 55 RGYFPPSHPSQDSGLLISRSHEPNELKKWITHENHRVRRMVP----- 96

a;Cross-references: GB:M25533; NID:9202126; PIDN:AAA4047.1; PID:g202127  
 c;Superfamily: Cysteine-rich secretory protein 1

QY 131 WDEVKQYTYPSPSECAPWC-----ERCSGPCTWVQIWTQWATNKIGAVNICKRM----- 184  
 Db 150 WFNEVH-----NPRCCGNHAYKHC-----CGHVQVWAKINTLVGGFSSRQDV 195  
 Qy 185 WGEVWENAVVFCNCNSPKGWI-----GE-----TPYKNGRPCSCCPPSPGCSR 230  
 Db 196 WGRHRHN-----FVCHNPQGMWVFTARGOLYAMPAFTWASGDGR-----CSCPAN-APACY 251  
 Qy 231 NNLCYREETTPKPERDEMHEVETAPIPERIPEN-----VET-----PTTICEPDEPEAEGADN 296  
 Db 252 OGLCKWPKNPKEAPTITE-----SITTSITEE-----PTTICEPDEPEAEGADN 296

RESULT 8  
 JE0204 testicular protein Tpx-1 - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Accession: JE0204  
 C;Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #Text\_change 21-Jul-2000  
 A;Position: 1-243 Res. Commun. 248, 140-146, 1998  
 R;Map position: 17  
 A;Title: Molecular cloning of the rat tpx-1 responsible for the interaction between sperm and biotin. Biophys. Res. Commun. 248, 140-146, 1998  
 R;Mada, T.; Sakashita, M.; Ohba, Y.; Nakanishi, Y.

QY 56 KEILMLMHNLKGQVOPQASMEYMMWDELEKSAAWASOCGCIWENGPT-----SLIVSIGO 112  
 Db 38 QREIVWKHNLRSVPICGSDIKMENWSQATNAQWANKTILEHSKDKRINKTRCE 97

QY 113 NLGAHWGRYRSP-----GFHQSWEVKDYYTYPSPSECNPWCPERCSPGPMCTHYQIWIW 167  
 Db 98 NL-----YMSTDPTLMSITVIOSWYNFNEDEFVYVGAK-----PNSAVG-----HTQLWV 142

QY 168 ATNUKIGCAVNICKRMWVGEWENAVVFCNCNSPKGWI-----GEAFYKNGRPCSCCPPS 224  
 Db 143 YSFHKICGIGIACPNO-----DNLYKFVYCHYCFCMGNNVMKKSTPYQOGTPCACP- 193

QY 225 YGSCRNLLC 234  
 Db 194 --NNCENGIC 201

RESULT 10  
 A44583 venom allergen antigen Ves f 5 - yellowjacket (Vespa flavopilosa)  
 C;Species: Vespa flavopilosa  
 C;Accession: A44583; B44522  
 R;Hoffman, D. R.  
 J;Allergy Clin. Immunol. 92, 707-716, 1993  
 A;Title: Allergens in hymenoptera venom XXV: the amino acid sequences of antigen 5 major;Reference number: A44583; MUID:9404416  
 A;Accession: A44583  
 A;Molecule type: protein  
 A;Residues: 1-204 <HOF>  
 C;Superfamily: yellowjacket venom allergen antigen 5

Query Match 9.7%; Score 268.5; DB 2; Length 243;  
 Best Local Similarity 31.7%; Pred. No. 2.3e-13;  
 Matches 60; Conservative 29; Mismatches 65; Indels 35; Gaps 8;

QY 56 KEILMLMHNLKGQVOPQASMEYMMWDELEKSAAWASOCGCIWENGPT-----SLIVSIGO 112  
 Db 38 QREIVWKHNLRSVPICGSDIKMENWSQATNAQWANKTILEHSKDKRINKTRCE 97

QY 113 NL-----GAHGRYKNSPGFHVQSWDEVKOYTYPSPSECNPWCERCSPPMCTHYQIWIW 167  
 Db 98 NLMSMHDPTSWRTV-----IQSWEEENENFVFGWAK-----PNSAVG-----HTQLWV 142

QY 168 ATTNKIGCAVNICKRMWVGEWENAVVFCNCNSPKGWI-----GEAFYKNGRPCSCCPPS 224  
 Db 143 YSSFRVGGVYACPNQDTL-----KYFYVCHYCFCMGNNVMKKSTPYQHGTCASC-----193

226 GGSRNLLC 234  
 194 -NNCDNGLC 201

RESULT 9  
 ESULT 9  
 testis-specific protein - mouse  
 Species: Mus musculus (house mouse)  
 ;Accession: A33329  
 Kasahara, M., Gutknecht, J.; Brew, K.; Spurr, N.; Goodfellow, P.N.  
 Title: Cloning and mapping of a testis-specific gene with sequence similarity to a species 5, 527-534, 1989  
 Reference number: A33329; MUID:90129048  
 Status: preliminary  
 Molecule type: mRNA  
 Residues: 1-243 <KAS>

RESULT 11  
 B37330 venom allergen III - red imported fire ant  
 N;Alternate names: allergen Sol I III  
 C;Species: Solenopsis invicta (red imported fire ant)  
 C;Accession: A33329  
 C;Date: 01-Oct-1992 #sequence\_revision 01-Oct-1992 #text\_change 11-Jan-2000  
 Hoffmann, D.R.  
 J;Allergy Clin. Immunol. 91, 71-78 1993

QY 161 HYTOIWATNKIGCAVNICKRMWVGEWENAVVFCNCNSPKGWNIGARYK 213  
 Db 156 HYTOMWANTKEVGG-----SIKFOEKNHH-YLVCNWPSPNFQNBELEYQ 202

A; Reference number: A44582; MUID:93139387  
 A; Accession: C44582  
 A; Status: preliminary  
 A; Molecule type: protein  
 A; Residues: 1-212 <H03>  
 R; Hoffman, D.R.; Smith, A.M.; Schmidt, M.; Moffitt, J.E.; Guralnick, M.  
 J. Allergy Clin. Immunol. 85, 988-995, 1990  
 A; Title: Allergens in Hymenoptera venom. Comparison of venoms from two species of import  
 A; Reference number: A60727; MUID:90285439  
 A; Accession: B60727  
 A; Molecule type: protein  
 A; Residues: 1-3, 'X', 5-8, 'X', 10, 'X', 12-15, 'X', 17-18, 'X', 20-23 <H02>  
 C; Superfamily: Yellowjacket venom allergen antigen 5  
 C; Keywords: venom

Query Match 9.0%; Score 249.5; DB 2; Length 212;  
 Best Local Similarity 35.6%; Pred. No. 5.8e-12;  
 Matches 64; Conservative 24; Mismatches 59; Indels 33; Gaps 7;  
 QY 55 DKEEILMHNKLRGQV-----OPOASNMNEYMTWDDLELEKSAAWASOCITWEGH 102  
 Db 42 EKDIAVWKHNLQRQVASGKEMRGTINGPQPAPAKMNPNTLWDPELATINQRWANQCTFEHD 101  
 QY 103 PRSLL--VSIGONLGA--HWGIVRS-PGFHYQSWSWDEVKDY---TYPYPSCNCNPWPER 153  
 Db 102 AGRNVERFRAVGONIATSSSGRNKNSTPNEMILWVNEVKDFDNRWRWSSFPSPDDNLM-- 158  
 QY 154 CSGPMCPHYTOIWATNTKIGCAVNTRKMTWGEWENAVYFVCVNYSPKGWIGEAPYK 213  
 Db 159 ---KVHBYTOIWAKTSKISKCARIMKFEDPNW---TKHVLVCHNGPAGNVLGAPIYE 209

RESULT 12  
 C44583  
 venom allergen antigen Ves p 5 - western yellowjacket  
 C; Species: *Vespaula maculifrons* (western yellowjacket)  
 C; Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994  
 C; Accession: C44583; C44522  
 R; Hoffman, D.R.  
 J. Allergy Clin. Immunol. 92, 707-716, 1993  
 A; Title: Allergens in hymenoptera venom XXV: the amino acid sequences of antigen 5 molecule  
 A; Reference number: A44583; MUID:94044316  
 A; Accession: C44583  
 A; Status: preliminary  
 A; Molecule type: protein  
 A; Residues: 1-204 <H02>  
 C; Superfamily: Yellowjacket venom allergen antigen 5

Query Match 9.0%; Score 249.5; DB 2; Length 212;  
 Best Local Similarity 35.6%; Pred. No. 5.8e-12;  
 Matches 64; Conservative 24; Mismatches 59; Indels 33; Gaps 7;

antigen 5 - eastern yellowjacket  
 C; Species: *Vespaula maculifrons* (eastern yellowjacket)  
 C; Date: 06-Nov-1992 #sequence\_revision 06-Nov-1992 #text\_change 11-Jan-2000  
 C; Accession: B37329  
 R; Lu, G.; Villalba, M.; Coscia, M.R.; Hoffman, D.R.; King, T.P.  
 Submitted to the Protein Sequence Database, August 1992  
 A; Reference number: A37329  
 A; Accession: B37329  
 A; Status: preliminary  
 A; Molecule type: protein  
 A; Residues: 1-204 <H01>  
 C; Superfamily: Yellowjacket venom allergen antigen 5  
 Query Match 8.7%; Score 240; DB 2; Length 204;  
 Best Local Similarity 29.6%; Pred. No. 3e-11;  
 Matches 68; Conservative 27; Mismatches 85; Indels 50; Gaps 9;

QY 1 MSCVIGGVIPFLGLFLVGCGSQYLIPNVTLLELLSKYOHNEHSRVRRAIPREDKEEL 60  
 Db 6 IKCLKGKV----HTAC-KYSLKPNC-----GNKKVSVSYGLTKQEKDIL 45  
 QY 61 MHNLKLRGQV-----OPOASNMNEYMTWDDLELEKSAAWASOCITWEGHPTSLV 108  
 Db 46 KEHNDFRQKIRGLETGRNPGOPPAKMNKLWSDELAYIAQWANOQCOYHDTCDVA 105  
 QY 109 --SIGQNL--GAHWGRYPSGPFWQSWSWDEVKDYTYPSPECNPWCFCRCSGPMTHT 163  
 Db 106 KYQVGQVALGTSTAAVNDPVPVLUKMNDEDEVKDYNPKRKESNNEL-----KIGHT 158  
 QY 164 QIWATNTKIGCAVNTRKMTWGEWENAVYFVCVNYSPKGWIGEADYK 213  
 Db 159 QMWQWANTKEVGC-----SIKYQENWIKH-LYLVCNIGPSPGNFQNEELYQ 202

RESULT 14  
 S60684  
 cysteine-rich secretory protein 1 precursor - human  
 C; Species: Homo sapiens (man)  
 C; Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 20-Jun-2000  
 R; Kraetschmer, J.; Haendler, B.; Eberspaecher, U.; Roosterman, D.; Donner, P.; Schleifer, J.; Bichem, 236, 827-836, 1996  
 A; Title: The human cysteine-rich secretory protein (CRISP) family. Primary structure  
 A; Reference number: S60681; MUID:96270732  
 A; Accession: S60684  
 A; Status: preliminary  
 A; Molecule type: mRNA  
 A; Residues: 1-249 <KRA>  
 A; Cross references: EMBL:X95237; NID:91262814; PIDN:CAA64524.1; PID:g1262815  
 C; Superfamily: cysteine-rich secretory protein 1  
 F; 1-21/Domain: signal sequence #status predicted <SG>  
 F; 22-249/Product: cysteine-rich secretory protein 1 #status predicted <MAT>

Query Match 8.7%; Score 239.5; DB 2; Length 249;  
 Best Local Similarity 29.6%; Pred. No. 4.2e-11;  
 Matches 74; Conservative 36; Mismatches 81; Indels 59; Gaps 13;

QY 13 LIFLVCGSQYLIPIPNVTLLELLSKYOHNEHSRVRRAIPREDKEELMLHNLRLRGQVQ 72  
 Db 6 LILFLVAA---CILPMLS---MKKSARDQDNPLVTDLPV-QEEIVNIIHARRRVVP 57  
 QY 73 QASNMNEYMTWDDLELEKSAAWASOCITWEGHTSLLS--IGQNLCAHWGRYSPCFHV 128  
 Db 58 PASNLKMSWSEAAQARIFSKYCDMIESNLLERRLNPTEGENM-----HM 105  
 QY 129 QSWYDEVKDYTYPPSCNPWPERCS---GPMCT-----HYTIWATNTKIGCAV 177  
 Db 106 TS-----YPSWSSVIGVWYSESTSPKHGWTDDDITDHYTIWATSYLIGCAI 158  
 QY 178 NTCRKMVTWGEWENAVYFVCVNYSPKGWIGEAPYK 213  
 QY 155 GHYTQWQWANTKEIGC-----GSKIVIONWKKH-YLVCVNGPSNFGNELYQ 202

RESULT 13  
 B37329

Db 159 ASCRQQ-----GSPRLYVCHYCSEGNDPETKNEPYRTGVPCACP----SNCEDKLCT 208  
Qy 235 ---YREETY 240  
| : | : |  
Db 209 NPCIYYDEF 218

RESULT 15  
G44583  
venom allergen antigen Vesp c 5.01 - European hornet  
C:Species: *Vespa crabro* (European hornet)  
C:Date: 27-jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 11-Jan-2000  
C:Accession: G44583; G44522  
J. Allergy Clin. Immunol. 92: 707-716, 1993  
A:Title: Allergens in hymenoptera venom XXV: the amino acid sequences of antigen 5 molec  
A:Reference number: A44583; MUID:94044316  
A:Accession: G44583  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-202 <HOF>  
C:Superfamily: yellowjacket venom allergen antigen 5

Search completed: May 25, 2002, 14:30:09  
Job time: 3710 sec

Title: US-09-667-380a-2  
 Perfect score: 2759  
 Sequence: 1 MSCVLLGGVTPPLGLFLIVCGS.....SESLGTPRDGKAFRIFAVRQ 497

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 14:28:55 ; Search time 63.23 Seconds  
 (without alignments)  
 304.343 Million cell updates/sec

Scoring table: BLOSUM62

Searched: 105224 seqs, 38719550 residues

Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SwissProt\_40;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

Result No.	Score	Query Match	Length	DB ID	Description
1	355	12.9	266	1 GLIP_HUMAN	GLIP_HUMAN
2	301.5	10.9	245	1 AEGI_MOUSE	ID: GLIP_HUMAN STANDARD; PRT; 266 AA.
3	298	10.8	245	1 ORS3_HORSE	AC: P48060; Q15409;
4	297.5	10.8	245	1 TPL1_HUMAN	DT: 01-FEB-1996 (Rel. 33, Created)
5	281.1	10.2	243	1 YR81_CAEEL	DT: 01-NOV-1997 (Rel. 35, Last sequence update)
6	269	9.7	312	1 YR81_CAEEL	DT: 16-OCT-2001 (Rel. 40, Last annotation update)
7	268.5	9.7	244	1 TXP1_CAVPO	DE: Glialoma pathogenesis-related protein (GIIPR) (RTVP-1 protein).
8	267.5	9.7	243	1 TXP1_MOUSE	GN: GIIPR OR RTVP1.
9	257.5	9.3	242	1 HBL0_HELI0	RA: Homo sapiens (Human).
10	256	9.3	227	1 VA5_VESCR	RC: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
11	253	9.2	204	1 VA5_VESFL	OS: Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
12	244	8.8	204	1 VA5_VESPE	OC: NCBI_TaxID=9605;
13	243.5	8.8	234	1 VA5_SOILIN	RN: [1] SEQUENCE FROM N.A.
14	240	8.7	204	1 VA5_VESNC	RN: TISSUE=Glial tumor; MEDLINE=97128816; PubMed=8973356;
15	239.5	8.7	249	1 CFS1_HUMAN	RN: Rich T., Chen P., Furman F., Buynh N., Israel M.A.; "RTVP-1, a novel human gene with sequence similarity to genes of diverse species, is expressed in tumor cell lines of glial but not neuronal origin"; Gene 180:125-130(1996).
16	236	8.6	202	1 VA51_VESCR	RN: [2] SEQUENCE OF 11-266 FROM N.A.
17	238	8.6	204	1 VA5_VESGE	RX: MEDLINE=95331646; PubMed=607567;
18	238	8.6	205	1 VA5_VESVI	RA: Murphy E.V., Zhang Y., Zhu W., Biggs J.J.; "The human glioma pathogenesis-related protein is structurally related to plant pathogenesis-related proteins and its gene is expressed specifically in brain tumors."; Gene 159:131-135(1995).
19	237	8.6	202	1 VA52_VESCR	RN: [3] STRUCTURE BY NMR; MEDLINE=98151500; PubMed=9482873;
20	237	8.6	203	1 VA5_DOLLAR	RN: Szyperski T., Fernandez C., Wuthrich K.; "Structure comparison of human glioma pathogenesis related protein P14a indicates a functional link between the human immune system and a plant defense system"; J. Biol. Chem. 273:22662-22667(1998).
21	237	8.6	211	1 VA3_SOLRI	RN: PROC. Natl. Acad. Sci. U.S.A. 95:2262-2266(1998).
22	236.5	8.6	424	1 ASI_ANCCA	RN: -1. TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE HUMAN BRAIN TUMOR, GLIOBLASTOMA MULTIFORM/ASTROCYTOMA, BUT NEITHER IN NORMAL FETAL OR ADULT BRAIN TISSUE, NOR IN OTHER NERVOUS SYSTEM TUMORS.
23	235	8.5	202	1 VA5_VESMA	RN: CC: -1. SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TXP1; INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
24	235	8.5	205	1 VA5_POLFU	RN: CC: This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce.html">http://www.isb-sib.ch/announce.html</a> or send an email to license@isb-sib.ch).
25	233.5	8.5	205	1 VA5_POLME	RN: DR: EMBL; X91911; CAA63005.1; -;
26	232	8.4	173	1 PRL_MEDTR	RN: DR: HSSP; P04284; ICFE.
27	230	8.3	206	1 VA5_POLDO	RN: DR: HSSP; P04284; ICFE.
28	222.5	8.1	205	1 VA5_POLEX	RN: DR: HSSP; P04284; ICFE.
29	222.5	8.1	209	1 VA5_POLAN	RN: DR: HSSP; P04284; ICFE.
30	221.5	8.0	215	1 VA5_DOLMA	RN: DR: HSSP; P04284; ICFE.
31	219.5	8.0	183	1 CRVP_TRIMU	RN: DR: HSSP; P04284; ICFE.
32	217	7.9	227	1 VA52_DOLMA	RN: DR: HSSP; P04284; ICFE.

P35794 schizophyllus musculus  
 P03602 mus musculus  
 P08299 nicotiana tabacum  
 P11670 nicotiana tabacum  
 Q04108 lycopersicum esculentum  
 Q00008 zea mays (mexicanus)  
 P07053 nicotiana tabacum  
 P09042 nicotiana tabacum  
 P35795 schizophyllum commune  
 P35792 hordeum vulgare  
 Q05968 hordeum vulgare

## ALIGNMENTS

DR	MIM: 602692; -.	CC	-1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPLX1;	
DR	InterPro: IPR001283; SCP.	CC	InSECTS A3/AG5; FUNG SC7/SC14 AND PLANTS PR-1.	
DR	Pfam: PF00188; SCP; 1.	CC	-----	
DR	PRINTS; PRO00837; V5TP1LIKE.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
DR	PROSITE; PS01009; SCP-AG5_PRL_SC7_1; 1.	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
DR	PROSITE; PS01010; SCP-AG5_PRL_SC7_2; 1.	CC	use by non-profit institutions as long as its content is in no way	
FT	D > N (IN REF. 2).	CC	modified and this statement is not removed. Usage by and for commercial	
FT	KRDQKRYSYVPPWPIPRNRTSLEFLVNSVILS	CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ) or send an email to license@isb-sib.ch).	
FT	LITIIVWLVKPVNLID->DSETRSNVNTMLYIRLAHS	CC	-----	
SQ	SEQUENCE 266 AA; 30342 MW; C0437808:EFB257A CRC64;	CC	-----	
DR	EMBL; X04643; CA28304_1; -.	DR	DR	
DR	Best Local Similarity 12.9%; Score 355; DB 1; Length 266;	DR	EMBL; M31173; AB59716_1; -.	
Matches 82; Conservative 33; Mismatches 75; Indels 40; Gaps 12;	DR	PIR; A24609; A24609.		
Db	5 LATIATMWSPFVSNYSH--ANTIPDIENEDFKDCVHLINKFRSEVEKPTASDMLYMWD 61	DR	PIR; A40918; A40918.	
QY	25 LPNVILLELLSKOHNEHSRVRRAIPED-KREBILMLHANKRGQYOPQASMEYMW 83	DR	InterPro: IPR001283; SCP.	
Db	84 DELKSAAWASOCQIWEHQG---PTSL---LVSIGOMLAGHGRGYNSPGFHVQS---W 131	DR	PRINTS; PRO00188; SCP; 1.	
Db	62 PALAQIKAWASACQFSINTRLAEPPHKHPNFTSGENI---W-TGSVPFESVSAITNW 117	DR	PROSITE; PS01009; SCP-AG5_PRL_SC7_1; 1.	
QY	132 YDEWKDYTVPPSPCNPWCPEPAGSMOTHYIQIWIWATTNKGCavitCRMFWGEWVE 191	DR	PROSITE; PS01010; SCP-AG5_PRL_SC7_2; 1.	
Db	118 YDETKDVF-----KTRICKVCGHYIQVWVADSYKVGCAVQFCPKVSGF-DALS 166	DR	SPERM; Glycoprotein; Signal.	
QY	192 NAVYFVCNYSPKGN--WIGEARYKNRGPSCPSPGGSCRNLYCREE 238	FT	CHIRL 20 246 SPERM-COATING GLYCOPROTEIN.	
Db	167 NGAHFICNYGPQGNPWTW---PYKRGATCSACPN--DKCLDNLCVNRO 210	FT	MOD_RES 20 20 BLOCKED (POTENTIAL).	
RESULT 2	AGS_RAT	STANDARD; PRT; 246 AA.	FT	CARBONID 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
ID	P12020;	DB	CARBONID 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).	
AC	01-OCT-1989 (Rel. 12, Created)	DB	CARBONID 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).	
DT	01-OCT-1989 (Rel. 12, Last sequence update)	DB	CARBONYD 213 213 N-LINKED (GLCNAC. . .) (POTENTIAL).	
DE	Sperm-coating glycoprotein precursor (SCP) (Acidic epididymal glycoprotein) (Protein D) (Protein E) (Protein IV) (Sialoprotein)	SO	SEQUENCE 246 AA; 27847 MW; 585207C7CF9D1 CRC64;	
DE	(32 kDa epididymal protein).	QY	11 LGILETLVCGSQGVLLPNTLLELLSKQHN ESHSRVRAIPREDKEELMLHNLKLRQ 69	
OS	Rattus norvegicus (Rat).	Db	5 LVLVLFLAA----VLP-PSILLDTEMDRDLNLSTIKLW---QEEIINKHNLRLRT 54	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Rattus.	QY	70 VQPOQASNMYMTWDELERSAAWASOCQIWEHQG---TSLSLISIGOML---GAWGRY 121	
OX	NCBI_TAXID=10116;	Db	55 VSPSGSDLRLVENDHDATVNAQKWARCINHNSPQLQRTTILKGGENLFMANYPASWSV 114	
RN	SEQUENCE FROM N.A.	QY	122 RSPGFHVOSSWYDVKDYTVPPSPCNPWCPEPAGSMOTHYIQIWIWATTNKGCavitCR 181	
RC	TISSUE=Epithelium;	Db	115 -----TQDWYBSDLRFGC-----PKV-GVKGHYYQVWNSTFLVAGCVAECP 160	
RX	MEDLINE=87053995; PubMed=3780731;	QY	182 KMWVGEWVANAVYFVNYSPPGNNWGE--APYKNGRGPCEBCPPPSYGGSCRNLYCREE 239	
RA	Brooks D.E., Means A.R., Wright E.J., Singh S.P., Tiver K.K.; "Molecular cloning of the cDNA for androgen-dependent sperm-coating glycoproteins secreted by the rat epididymis.", Eur. J. Biochem. 161:13-18(1986). [2]	Db	161 DPL-----KPYFVCHICPGFNNYVGRLYSPPTGECPCDSCP---GNCEDGICTNSE 209	
RX	SEQUENCE FROM N.A.	QY	240 YPKPENDEMNEVETAPIEPERNHVMQPRMPTKPKKTSAVNINTQVVRCDTK-MKDR 238	
RA	MEDLINE=89039913; PubMed=2460753;	Db	210 Y-----EDN-----SNCGDQKKMVSCDDPLIKECC 235	
RT	"Molecular cloning of complementary deoxyribonucleic acid for an androgen-regulated epididymal protein: sequence homology with metalloproteins."	QY	299 KGS 301	
RT	Mol. Endocrinol. 2:99-100(1988).	Db	236 RAS 238	
CC	-1- FUNCTION: THIS PROTEIN IS SUPPOSED TO HELP SPERMATOZOA UNDERGO FUNCTIONAL MATURATION WHILE THEY MOVE FROM THE TESTIS TO THE DUCTUS DEFERENS.	RESULT 3	-----	
CC	-1- INDUCTION: by androgens.	ID	AEG1_MOUSE	
CC	-1- MISCELLANEOUS: SCP IS AN ANDROGEN-DEPENDENT PROTEIN, WHICH IS SECRETED BY THE EPIDIDYMAL EPITHELIUM AND THEN BECOMES ASSOCIATED WITH THE SPERM SURFACE. TWO MAJOR VARIANT PROTEIN D AND E DIFFER FROM EACH OTHER BY THEIR CARBOHYDRATE SIDE CHAINS.	AC	003401; STANDARD; PRT; 244 AA.	
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	DT	01-OCT-1993 (Rel. 27, Created)	
CC		DT	01-OCT-1993 (Rel. 27, Last sequence update)	
CC		DT	01-MAR-2002 (Rel. 41, Last annotation update)	
CC		DE	Sperm-coating glycoprotein 1 precursor (SCP 1) (Acidic epididymal glycoprotein 1) (Cysteine-rich secretory protein-1) (Crisp-1)	
CC		DE	AEG1 OR AEG-1.	
CC		OS	Mus musculus (Mouse).	



QY 165 IWWATNNKIGCAVNTCRMTWGVENAVYFVCNYSPKGWIGE-APYKNRGPSECP 222  
 :|| ::|| :|| :|| ;|| || || || || || || || || || || || || || || ||  
 Db 142 VWWYSSTRVGCCAYICPKQ--GTL--KYYVVCQCPAGNVNKINTPYEQTCARCP 195  
 QY 223 PSYGGSCRNNIYCREEYTYPKPEDENEVEAIPFEEENHVLQPRVMRPTKKKSAVN 282  
 :|| || || || || || || || || || || || || || || || || || || || || ||  
 Db 196 ---GNCDNGLC-----TNSC 208  
 QY 283 YMMQVYVRCDT-----RMRDRKGSTCNRRQC 308  
 :|| || || || || || || || || || || || || || || || || || || || ||  
 Db 209 YEDLVSNCDSKKKIACCEHELLIKENCK-TWC--OC 240  
 RESULT 5  
 CRIS3\_HUMAN STANDARD; PRT; 245 AA.  
 ID CRIS3\_HUMAN STANDARD;  
 AC P54108; Q15512;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cysteine-rich secretory protein-3 precursor (CRISP-3) (SGP28 protein).  
 GN CRIS3  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SQUENCE FROM N.A.  
 RX MEDLINE:96270732; PubMed=86655901;  
 RA Kretzschmar J., Haenler B., Eberspaecher U., Roostermann D.,  
 RA Donner P., Schleuning W.-D.;  
 RT "The human cysteine-rich secretory protein (CRISP) family. Primary  
 structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3.",  
 Eur. J. Biochem. 236:827-836(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE:96166934; PubMed=8601434;  
 RA Kjeldsen L., Cowland J.B., Johnson A.H., Borregaard N.;  
 RT "SGP28, a novel matrix glycoprotein in specific granules of human  
 neutrophils with similarity to a human testis-specific gene product  
 and a rodent sperm-coating glycoprotein.",  
 J. Cell. Sci. 109:246-250(1996).  
 RL FBBs Lett. 380:246-250(1996).  
 CC -.1- SUBCELLULAR LOCATION: SECRETED; IN NEUTROPHILS,  
 SPECIFIC GRANULES.  
 CC -.1- TISSUE SPECIFICITY: SALIVARY GLAND, PANCREAS AND PROSTATE >  
 CC EPIDIDYMIS, OVARY, THYMUS AND COLON  
 CC -.1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;  
 CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 ----  
 RESULT 6  
 TPX1\_HUMAN STANDARD; PRT; 243 AA.  
 ID TPX1\_HUMAN STANDARD;  
 AC P16567;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Testis-specific protein TPX-1 precursor (Cysteine-rich secretory  
 protein-2) (crisp-2).  
 RN TPX1  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OX NCBI\_TAXID=9606;  
 RN [1]  
 RP SQUENCE FROM N.A.  
 RC TISSUE="testis";  
 RX MEDLINE:90123048; PubMed=2613236;  
 RA Kasahara M., Gutknecht J., Brew K., Spurr N., Goodfellow P.N.,  
 RT "Cloning and mapping of a testis-specific gene with sequence  
 similarity to a sperm-coating glycoprotein gene.",  
 Genomics 5:527-534(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE="testis";  
 RX MEDLINE:96270732; PubMed=86655901;  
 RA Kraetzschmar J., Haenler B., Eberspaecher U., Roostermann D.,  
 RA Donner P., Schleuning W.-D.;  
 RT "The human cysteine-rich secretory protein (CRISP) family. Primary  
 structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3.",  
 Eur. J. Biochem. 236:827-836(1996).  
 RL FBBs Lett. 380:246-250(1996).  
 CC -.1- SUBCELLULAR LOCATION: TESTIS AND EPIDIDYMIS.  
 CC -.1- TISSUE SPECIFICITY: TESTIS AND EPIDIDYMIS.  
 CC -.1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;  
 CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 ----



CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;  
 CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC used by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).  
 CC

CC EMBL; U35712; AAC52616.1; -.  
 DR HSSP; P04284; ICFP.  
 DR InterPro; IPR001283; SCP.  
 DR Pfam; PF00188; SCP; 1.  
 DR PRINTS; PRO0837; V5TPX1LIKE.  
 DR SMART; SM0198; SCP; 1.  
 DR PROSITE; PS01009; SCP\_AG5\_PR1\_SC7\_1; 1.  
 DR PROSITE; PS01010; SCP\_AG5\_PR1\_SC7\_2; 1.  
 DR KW PROSITE; PS01009; SCP\_AG5\_PR1\_SC7\_1; 1.  
 DR PROSITE; PS01010; SCP\_AG5\_PR1\_SC7\_2; 1.  
 FT SIGNAL; 1; 21  
 FT CHAIN; 22; 244 AA; 27248 MW; 58D8DE6ECE1A12 CRC64;  
 SQ SEQUENCE 244 AA; 27248 MW; 58D8DE6ECE1A12 CRC64;

Query Match 9.7%; Score 268.5; DB 1; Length 244;  
 Best Local Similarity 28.8%; Pred. No. 2.4e-14; Indels 67; Gaps 12;  
 Matches 74; Conservative 29; Mismatches 87; Indels 67; Gaps 12;

QY 56 KERILMLHNLKLRQVQPOASNMEMYTWDDELEKSAAWASOCIWENG---PSILSVI 110  
 Db 38 QREINLNKHNLKRKVTPAPSNMNLKNEWSREAVNAQWANKRCTLWHSNPDPRKTS---TRC 95  
 QY 111 GONL---GAHWGRYRSPGHVQWYDEVRYDTYPP--PSECNPCCPERCSGPACTHYIQ 164  
 Db 96 GENLYMSSDPSWSD-----AIQSWFDESQDFTEFGVKFRSHN-----AVVGHVHQ 140  
 QY 165 IWTNKKIGAVNCTKRMVWGEWENAVYFVNYSYSPGN--WIGEAPYKNGRPSCP 222  
 Db 141 LWVWSYSLVCGIAYCPNQD-----SLRVVYVCQCPAGNNVTKTNPYKOGIPASCAP 194  
 QY 223 PSYGSSCRNLICRYETTYPPKPEDEMNVETAPILEENHVWHLQPRVMPRKPKRTSAVN 282  
 Db 195 ---GHCENGLCCTNSCEY----EDLJSNCESL-----KNTAGCE 225  
 QY 283 YMTQWVRCDTKMDRCK 299  
 Db 226 HQLVLEKC--KATCRCE 240

RESULT 9  
 TPX1\_MOUSE STANDARD PRT; 243 AA.

AC P16563;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DE Testis-specific protein TPX-1 precursor.  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Buteraria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=20129048; PubMed=2613236;  
 RA Kasahara M., Gutknecht J., Brew K., Spurr N., Goodfellow P.N.;  
 RT "Cloning and mapping of a testis specific gene with sequence  
 similarity to a sperm coating glycoprotein gene.";  
 RL Genomics 5:527-534(1989).  
 CC -!- SUBCELLULAR LOCATION: Secreted (probable).  
 CC -!- TISSUE SPECIFICITY: TESTIS.

CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;  
 CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC used by non-profit institutions. There are no restrictions on its  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).  
 CC or send an email to license@ib-sib.ch).

CC EMBL; U35712; AAC52616.1; -.  
 DR HSSP; P04284; ICFP.  
 DR InterPro; IPR001283; SCP.  
 DR Pfam; PF00188; SCP; 1.  
 DR PRINTS; PRO0837; V5TPX1LIKE.  
 DR SMART; SM0198; SCP; 1.  
 DR PROSITE; PS01009; SCP\_AG5\_PR1\_SC7\_1; 1.  
 DR PROSITE; PS01010; SCP\_AG5\_PR1\_SC7\_2; 1.  
 DR KW PROSITE; PS01009; SCP\_AG5\_PR1\_SC7\_1; 1.  
 DR PROSITE; PS01010; SCP\_AG5\_PR1\_SC7\_2; 1.  
 FT SIGNAL; 1; 22  
 FT CHAIN; 23; 243 AA; 27605 MW; 6E707F569ACAA244 CRC64;

Query Match 9.7%; Score 267.5; DB 1; Length 243;  
 Best Local Similarity 33.2%; Pred. No. 2.8e-14; Indels 37; Gaps 9;  
 Matches 63; Conservative 28; Mismatches 62; Indels 37; Gaps 9;

QY 56 KEITMLHNLKLRQVQPOASNMEMYTWDDELEKSAAWASOCIWENG---PSILSVI 110  
 Db 38 QREINLNKHNLKRKVTPAPSNMNLKNEWSREAVNAQWANKRCTLWHSNPDPRKTS---TRC 95  
 QY 113 NLCAHWGRYRSP---GPHVWYDEVRYDTYPPSECNPCCPERCSGPACTHYIQ 167  
 Db 98 NL-----YMSDPTPLSTVQSQWYNEQEDVYVGAK---PNSAVG---HYTOLW 142  
 QY 168 ATWNKKIGAVNCTKRMVWGEWENAVYFVNYSYSPGN--WIGEAPYKNGRPSCP 224  
 Db 143 YSKFKIGGIAVCPNQ-----DNKIFPVCHYCPMGNNVMKKSTPYQOGTPCASP-- 193  
 QY 225 YGGSCRNLICRYETTYPPKPEDEMNVETAPILEENHVWHLQPRVMPRKPKRTSAVN 234  
 Db 194 --NNCENGGLC 201

RESULT 10  
 HELO\_HELHO STANDARD PRT; 242 AA.

AC P01055;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DE Helotherrmine precursor (HLPx).  
 OS Heloderma horridum horridum (Mexican beaded lizard).  
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Helodermatidae;  
 OC NCBI\_TaxID=8552;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Sialivary gland;  
 RX MEDLINE=95375162; PubMed=7647234;  
 RA Morrisette J., Kraetzschmar J., Haendler B., El-Hayek R.,  
 RA Mochca Morales J., Martin B.M., Patel J.R., Moss R.L.,  
 RA Schleuning W.-D., Coronado R., Possani L.D.,  
 RT "Primary structure and properties of helotherrmine, a peptide toxin  
 that blocksryanodine receptors.,";  
 RL Biophys. J. 68:2280-2288(1995).  
 RN [2]  
 RP SEQUENCE OF 20-39, AND CHARACTERIZATION.







Tue May 28 16:10:05 2002

us-09-667-380a-2.rsp

Page 10

SEARCHED NO: 2  
AC NO: AAYA 1738, Database: A-Genealogy, 03-2802  
<sup>Page 6</sup>

to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the polynucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH0166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH29446 to AAH25893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

CC  
XX  
SQ Sequence 371 AA;

Quarry Match 75.1%; Score 2071; DB 22; Length 371;

Best Local Similarity 99.7%; Pred No. 2.7e-169; Mismatches 0; Indels 0; Gaps 0;

QY  
1 MSCVLGGVPIGLLFLMCGSGQYLLPNTTLEELLSKQHNEHSRVRRAI PREDKEEIL 60

Db  
1 macvlggvpplglifvqsgsqy1pntvltleelisykyqnhesnrvraibleukeell 60  
61 MTHNKT RGWQPOASNNYMWTDDELEKSAAAWASQCSCTWEHGPSTSLVSGIONLGAHWGR 120

Db	121	yrspqfhvqsrydevkdjtypypscnppwcpcasgpmcthytqvawattnkigcavnlc	160
Q <sub>11</sub>	161	prkmwmcgevwewavfyvnyNSPKGNWIGEAPYNGRPGCSCCPPSGSSCRNNLGYREETY	240

Db 181 rkantwgeewenavycvncyspkgnawigaeapkyngprcsecppsgscrnlycreety 240

QY  
241 TPKPETDEMNVEVETAPIPEENHMLQPRVMRPTPKKTSAVNYMTQVRCDTKMDKRCKG 300

Db 241 tpkpetdamnevetapipenehnwvlqprvrptpkkttsavnytqvvcatkmkarcky 300  
Cys 301 smcpnwpccpccfcnnhakrtgslyeeesssicraahhyiiddkgglyvditrngkypffv 360

QY 361 KSERHCVQL 370

Db 361 kserqvgvql 370

RESULT 8  
AAY1738

ID AAY41738 standard; protein; 500 AA.  
XX  
AC AAV41738.

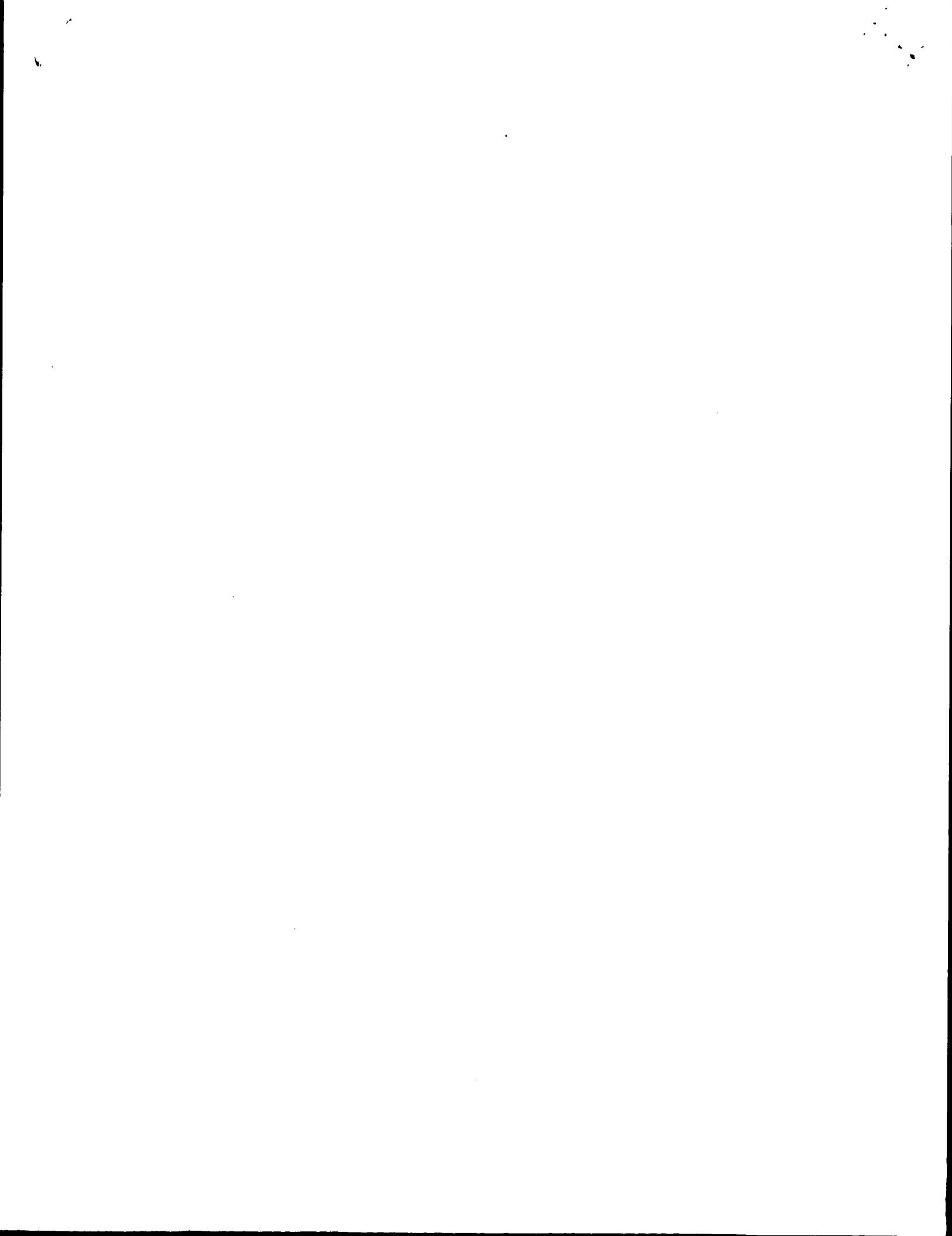
XX  
XX  
DT  
07-DEC-1999  
(first entry)

XX  
DE Human PRO541 protein sequence.  
XX

**KW** Human; PRO; EST; expressed sequence tag; PCR Primer; hybridisation; probe; blood coagulation disorder; cancer; cellular adhesion disorder; protein; transmembrane protein.

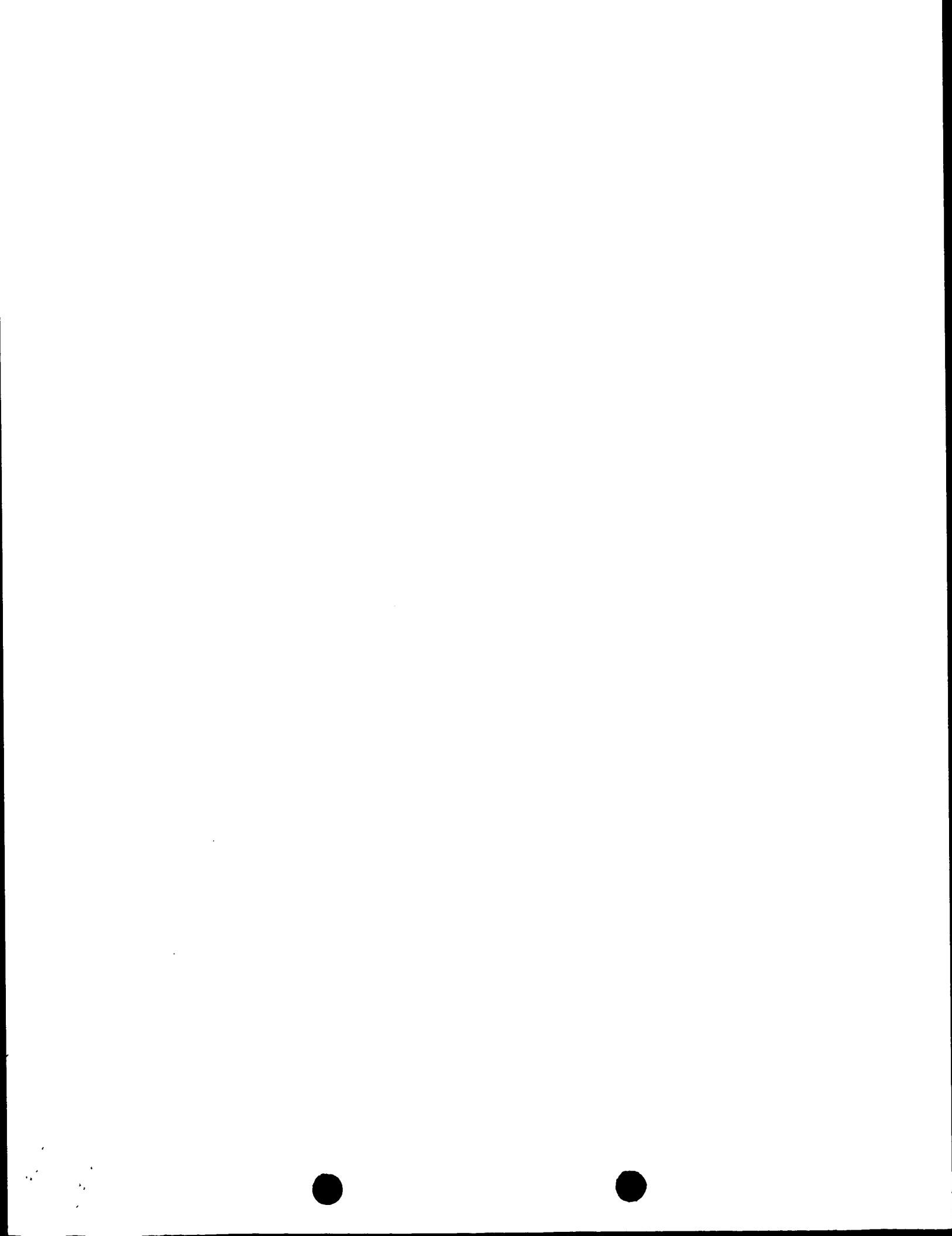
**XX** **XX** **OS** **Homo sapiens.**

XX  
PN  
WO9946281-A2.



SEA ID NO: 1  
AC No: AAS6084  
Database: N-Geneseg-032802

PI Zambrowicz B, Sands AT;  
XX  
DR WPI; 2001-266064/27.  
XX P-PSDB; AAB74446.  
XX  
XX Novel isolated human protease inhibitor-like polynucleotide useful in  
XX therapeutic, diagnostic and pharmacogenetic applications -  
PS Disclosure; Page 28-29; 29pp; English.  
XX  
CC The present invention provides the protein and coding sequences of the  
CC novel human protease-inhibitor like protein NHP. This shows homology to  
mammalian trypsin inhibitors. The sequences are useful in disease  
CC diagnosis and treatment, particularly of diseases associated with signifi-  
CC cant transduction. The present sequence is one version of the NHP cDNA.  
XX  
SQ Sequence 2272 BP; 532 A; 606 C; 649 G; 484 T; 1 other;



Page A

XX  
PR  
XX  
PA  
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI  
Lillie J., Brown JL, Bolt A, van Huffel C;  
XX  
DR  
WPI; 2001-602933/68.

XX  
PT  
Novel nucleic acid, used as a marker to determine the effectiveness of using TAXOL to treat cancer cell growth in individuals -  
XX  
PS  
Claim 1; Page 375-376; 527pp; English.

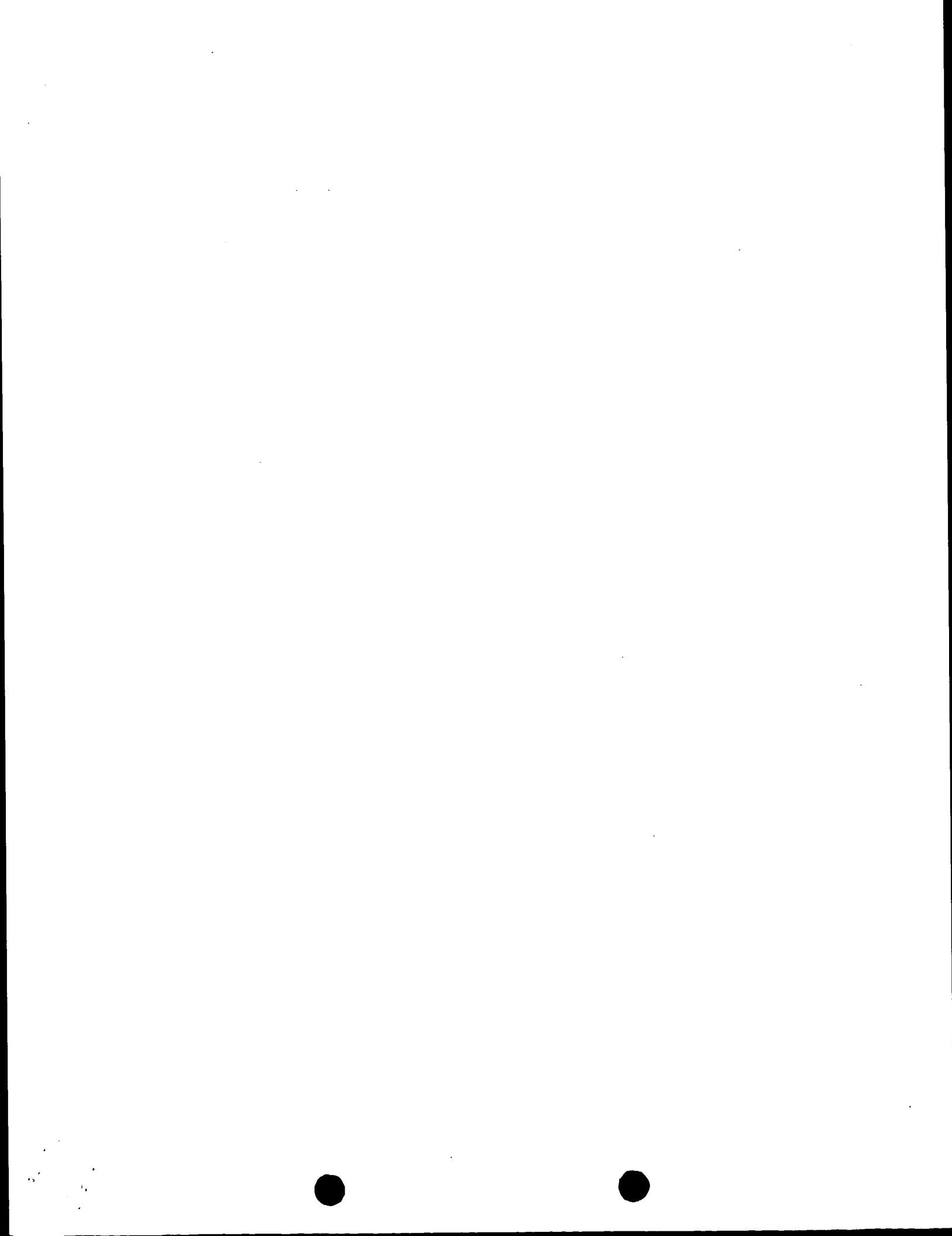
CC The invention relates to 1046 novel nucleic acids which are used as markers for determining the sensitivity of a cancer cell to the anticancer agent TAXOL. Cancer cells can be treated with TAXOL when they are shown to express one of the 242 sensitivity markers or the cells are shown not to express one of the 804 resistance markers. The methods can be used to determine the effectiveness of TAXOL in the treatment of cancer cell growth in an individual. The markers can be used as targets in developing anti-cancer agents such as chemotherapeutic compounds. The markers can also be used as targets in developing treatments for cancer, particularly those cancers which display resistance to agents and exhibit expression of the markers. The anticancer agents developed by the novel method can be used to treat cancer. Probes based on the markers can be used to detect transcripts genomic sequences corresponding to the markers. In the identification of cells or tissues which mis-express the protein, cancers which may be targeted include carcinoma (e.g., squamous cell carcinoma), sarcoma (e.g., fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia), lymphoma, plasmacytoma, reticular cell sarcoma, Hodgkin's disease and tumours (e.g. glioma). The present sequence is one of the 1046 novel cancer cell markers.

卷之三

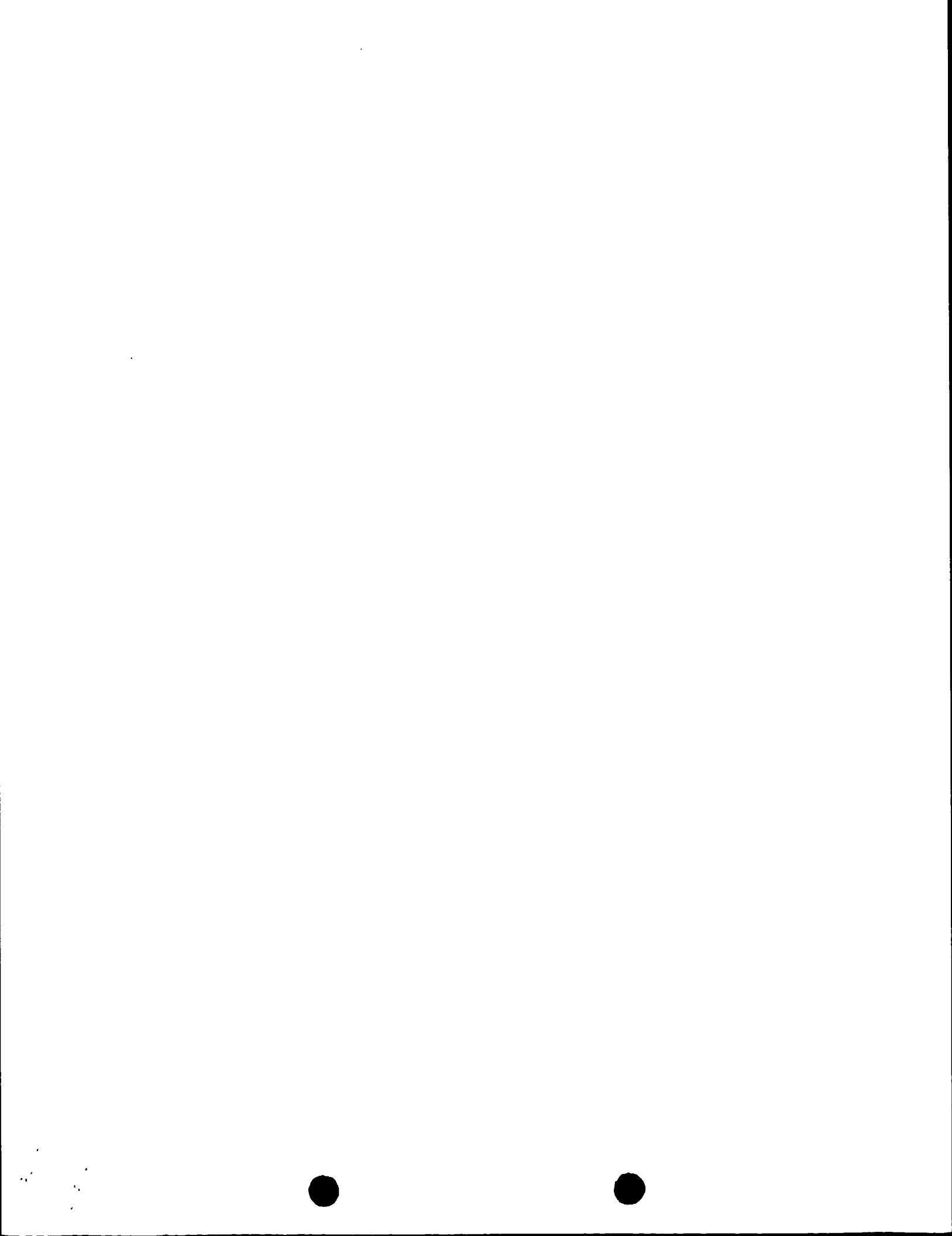
卷之三

Query Match 99.8%; Score 1487.8; DB 22; Length 4877;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1489; Conservative 0; Mismatches 2; Indels 0; Gaps

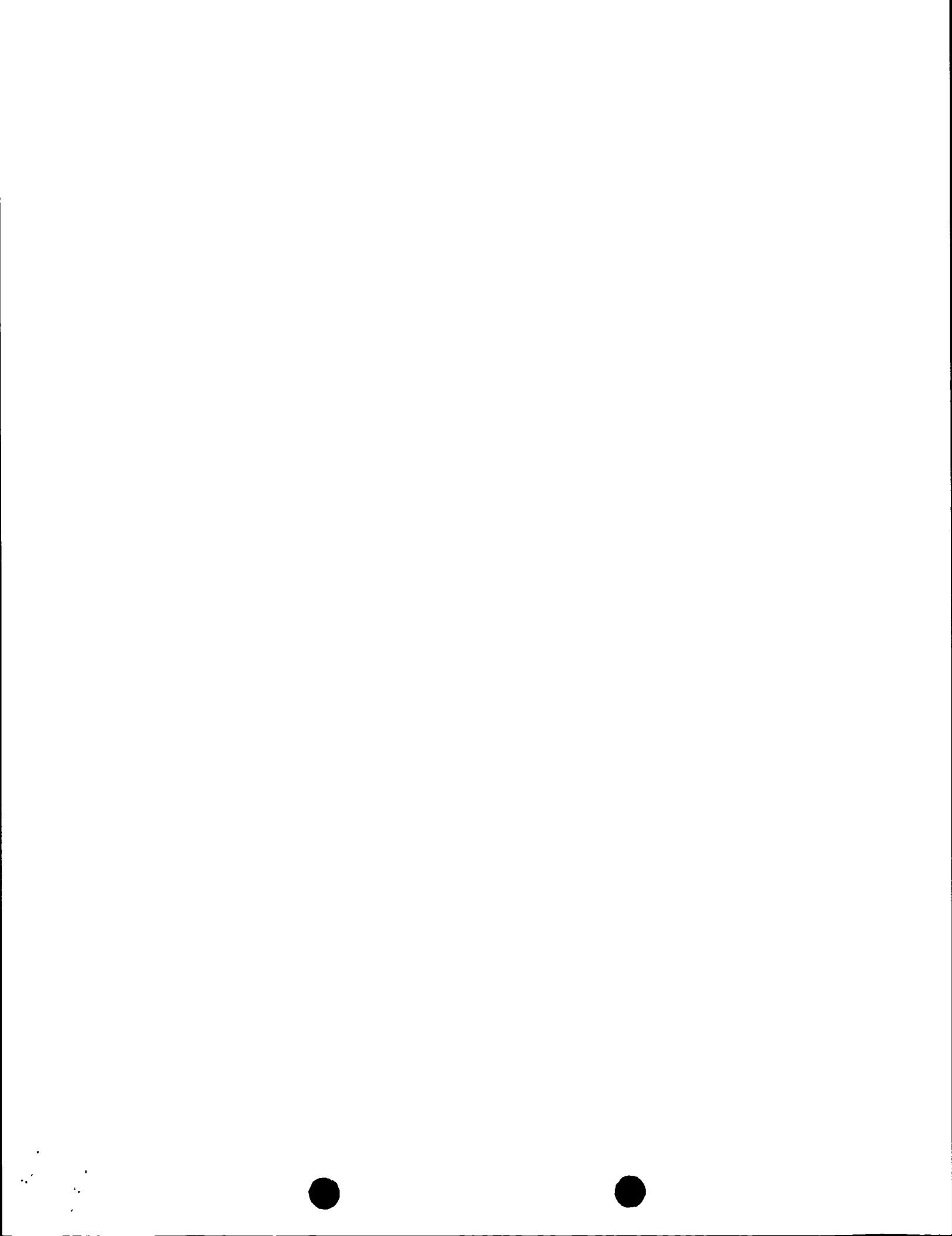
			0;
Db	3637	CTGGATGACAGGGAGGCCTGGATATCACCCAGGACGGGAGGTTCCCTTCCTG	3578
QY	1081	aagtctcgagaaacacggcgctcgatccctcgcaataacaaccccttcagatctatcg	1140
Db	3577	AATGTCGAGACACCGCGTGTGAGTCCTCAGCAAAATACACACCTTCAGTCATTAG	3518
QY	1141	gtgtcaaaaggaaatggactacacgacggattggactatacagcggtgtcagtggtgcgg	12000
Db	3517	GTTGTCAGAAAGTGAAAGTGACGATTTGGACTCTACAGCAGGTGTCAGCTGCGCG	3458
QY	1201	tttgaaaaggccggcaactcaactggcccaaagaatccatgttcccgacactgccaagacaa	12600
Db	3457	TTTGAAAGGCCAGCACTCAAGGCCAAGAATCCTGTCGGCACHTGCCAAGACCAA	3398
QY	1261	ctttcctactgtggctcggtttttggaaaccacatcatgtcgatatactcaagcatctgc	13200
Db	3397	CCTTCCTACTCTGCTGGGTTGGHACCAACATCTAGCAGATACCTCTAACACATCGC	3338
QY	1321	aagacgtcggtggcacgggggtcatcagcacgagatgggggtggcgatgtgtatgg	13800
Db	3337	AAGACAGCCGCGTCACCGGGGAGTCATCAGCACAGAGTGGGGTGAAGTGGAGCTGATG	3278
QY	1381	cccggtggataaaaaggaaaggacctactgtggcgctcgatggaaatgggttcgtatggaa	14400
Db	3277	CCCGTGGATACAAAGAAGACCTTACGTTGGCTGCTCAGGAATGGAGTTCAGTCGAAGC	3218
QY	1441	ctggggactccctcgggatggaaaggcccttcggatctttgtgtcgatgg	1491
Db	3217	CTGGGGACTCTCTGGGATGGAAAGGCCCTCCGGATCTTGTGCTGAGGCG	3167



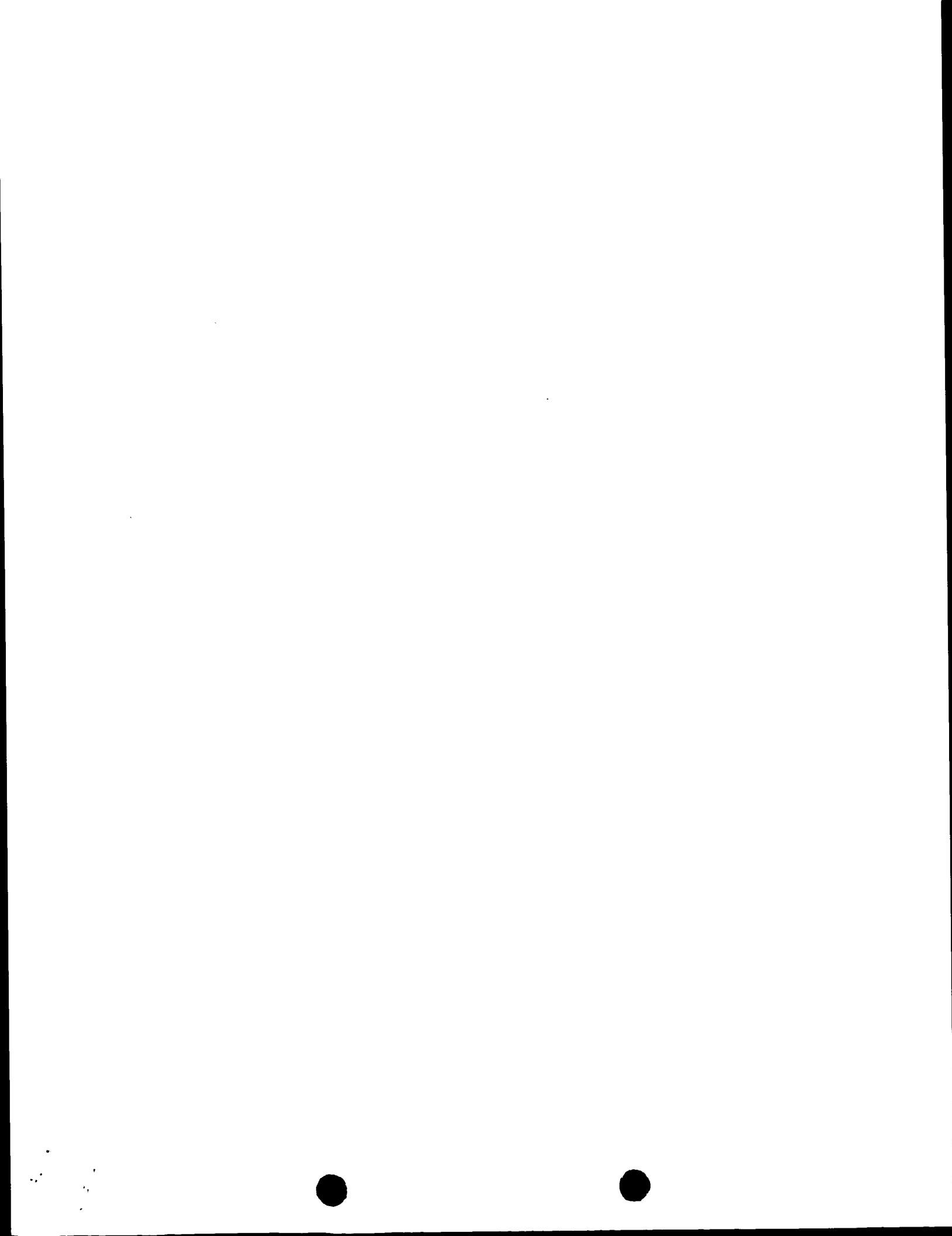












Db 61 MHNKLKGQVQFQASNNEMYMTWDELEKSAAWASOQIWEHGPTSLVISQNLGHWGR 120  
 QY 121 YRSPGFHVSQWSDDEVKDVTYPSECPNPWCBCRCSEPMCTHYTQIWATWKIGAVNTC 180  
 Db 121 YRSPGFHVSQWSDDEVKDVTYPSECPNPWCBCRCSEPMCTHYTQIWATWKIGAVNTC 180  
 QY 181 RKMTHGEVWENAVYFVCNTSPKGWIGEAPYKNGRPCSECPSYGCSCRNLCYREETY 240  
 Db 181 RKMTHGEVWENAVYFVCNTSPKGWIGEAPYKNGRPCSECPSYGCSCRNLCYREETY 240  
 QY 241 TPKEPEDEMNEVETAPIPEENHWLQPRVMRTPKPKTSAYNMTOVRCDTKMKDRKG 300  
 Db 241 TPKEPEDEMNEVETAPIPEENHWLQPRVMRTPKPKTSAYNMTOVRCDTKMKDRKG 300  
 QY 301 STCNRYOCPAGCLNHAKRIFGFLFESSSSICRAAHYGIIDDKGLVDTRNGKPFV 360  
 Db 301 STCNRYOCPAGCLNHAKRIFGFLFESSSSICRAAHYGIIDDKGLVDTRNGKPFV 360  
 QY 361 KSERHGVSLSKYKQSSFMVKVQDLCYTTAQLCPFKPATCPRHCPACKDE 420  
 Db 361 KSERHGVSLSKYKQSSFMVKVQDLCYTTAQLCPFKPATCPRHCPACKDE 420  
 QY 421 PSYWALVFGENIYANNSSTKTAVHAGVISENESGGDVMDVKKTYVESLRNQVSES 480  
 Db 421 PSYWALVFGENIYANNSSTKTAVHAGVISENESGGDVMDVKKTYVESLRNQVSES 480  
 QY 481 LGTPPDGKARIYARQ 497  
 Db 481 LGTPPDGKARIYARQ 497

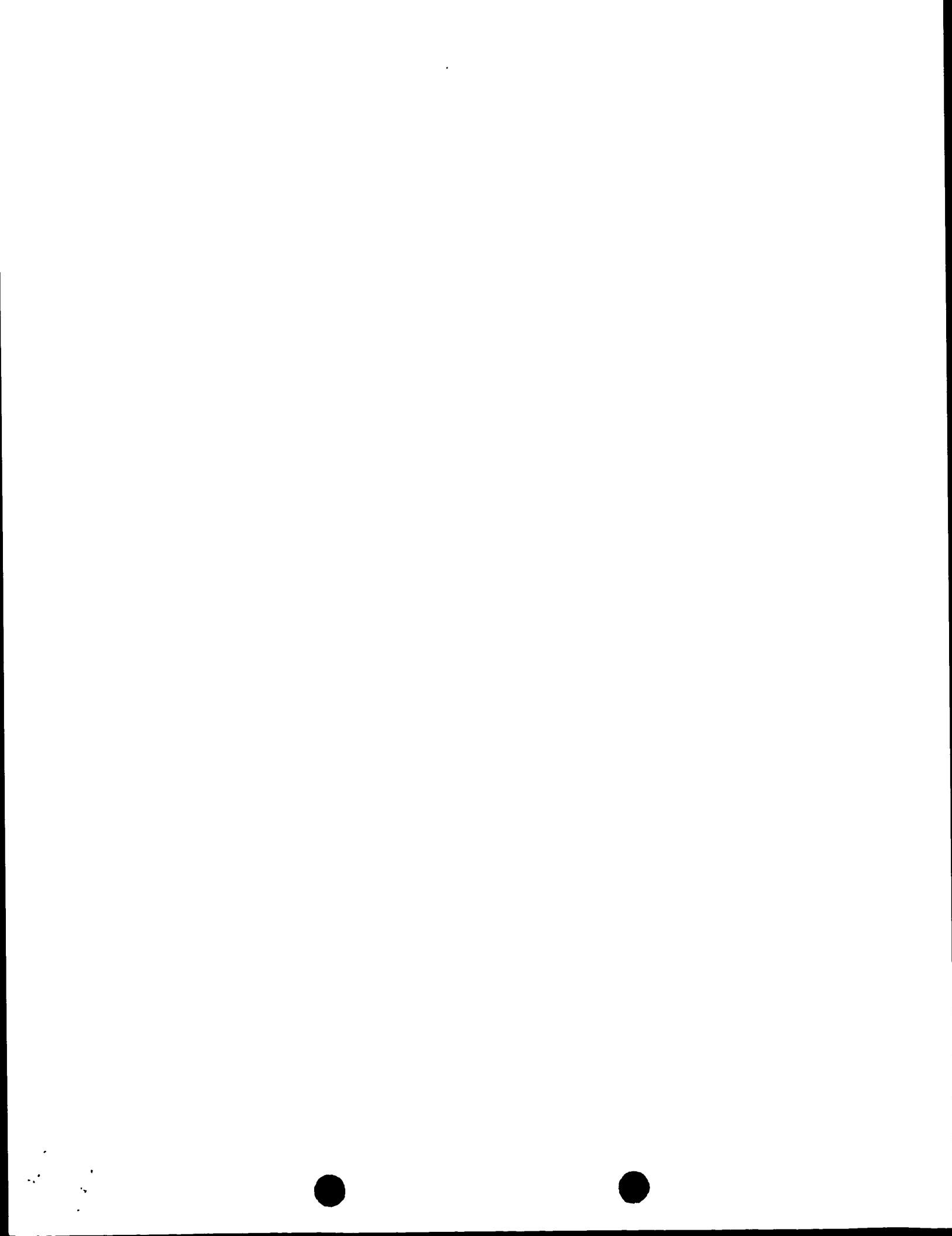
RESULT 2  
 ID 096K61 PRELIMINARY; PRT; 371 AA.  
 AC 096K61;  
 DT 01-DEC-2001 (TremblieL, 19, Created)  
 DT 01-DEC-2001 (TremblieL, 19, Last annotation update)  
 DE CDNA FJ14480, FIS, CLONE MAMM1002881, WEAKLY SIMILAR TO GLIOMA PATHOGENESIS-RELATED PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=100901  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=PANCREAS;  
 RR MEDLINE=2108560; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Futunishi T., Konno H., Adachi J., Fukuda S., Alzava K., Izawa M., Nishii K., Kiyosawa H., Kondo S., Yamakawa T., Saito T., Okazaki Y., Gojobori T., Poona H., Kasukawa T., Saito R., Kadota K., Matsuda H., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsumoto Y., Nikaido Y., Pesole G., Quackenbush J., Schriml L.M., Staubli R., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okada T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Masilama J., Mazzarelli M., Mombaert P., Nordeon P., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shimada Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynnshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S., Hayashizaki Y.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK027395; BAB55001.1; -  
 SQ SEQUENCE 371 AA; 42207 MW; 8D23FBEL14F55E85C CRC64;

Query Match Best Local Similarity 75.1%; Score 2071; DB 4; Length 371; Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSCVLGGVPLGLFLFGSGQYLLNNTLLELLSKYQHNTSHSRVRAIPREDKEIL 60  
 Db 1 MSCVLGGVPLGLFLFGSGQYLLNNTLLELLSKYQHNTSHSRVRAIPREDKEIL 60  
 QY 61 MHNKLKGQVQFQASNNEMYMTWDELEKSAAWASOQIWEHGPTSLVISQNLGHWGR 120  
 Db 61 MHNKLKGQVQFQASNNEMYMTWDELEKSAAWASOQIWEHGPTSLVISQNLGHWGR 120  
 QY 121 YRSPGFHVSQWSDDEVKDVTYPSECPNPWCBCRCSEPMCTHYTQIWATWKIGAVNTC 180  
 Db 121 YRSPGFHVSQWSDDEVKDVTYPSECPNPWCBCRCSEPMCTHYTQIWATWKIGAVNTC 180  
 QY 121 YRSPGFHVSQWSDDEVKDVTYPSECPNPWCBCRCSEPMCTHYTQIWATWKIGAVNTC 180

Query Match Best Local Similarity 79.6%; Score 1987.5; DB 11; Length 434; Matches 348; Conservative 37; Mismatches 49; Indels 3; Gaps 1;

QY 61 MHNKLKGQVQFQASNNEMYMTWDELEKSAAWASOQIWEHGPTSLVISQNLGHWGR 120



XX

PS

Claim 2; Page 27-28; 29PP; English.

FT

XX

CC

The present invention provides the protein and coding sequences of the novel human protease-inhibitor like protein NHP. This shows homology to mammalian trypsin inhibitors. The sequences are useful in disease diagnosis and treatment, particularly of diseases associated with signal transduction. The present sequence is the NHP protein.

Sequence 497 AA;

Query Match 100.0%; Score 2759; DB 22; Length 497;  
Best Local Similarity 100.0%; Pred. No. 3.6e-228; Mismatches 0; Indels 0; Gaps 0;  
Matches 497; Conservative 0; MisMatches 0; Domains 0;  
Protein /label= "Human mature novel trypsin inhibitor-like Protein, NOV-4b"  
Domain 195..206 /label= SCP\_domain

W0200162928-A2.

PN XX

PR XX

PD XX

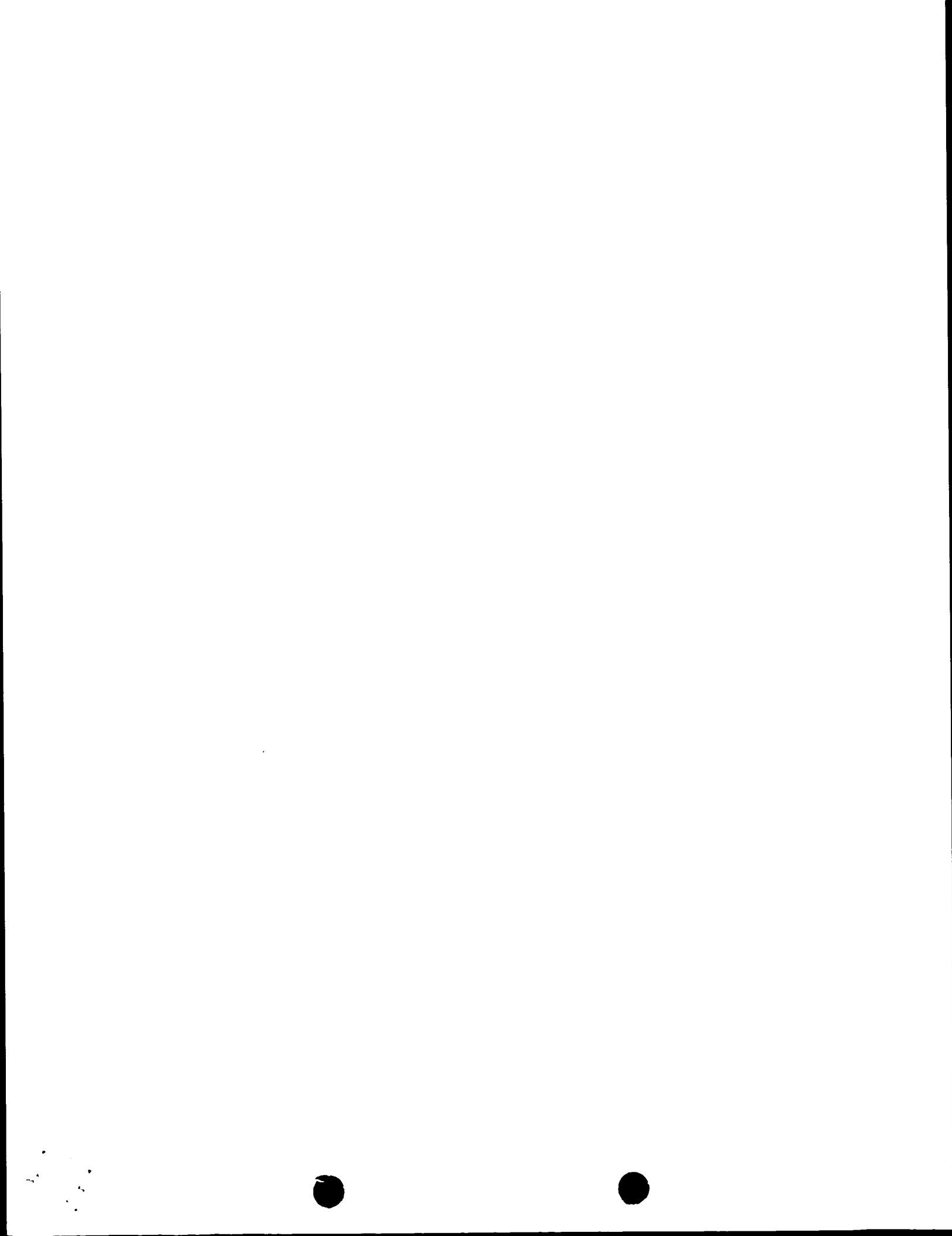
PP XX

PR XX

DR XX

N-PSD; AAD17764.

PT XX



Db 181 ||||||| rkmtywgewenavfyfcyspkgnwigeapvkgngpcsccepysqgscrnnlcyrety 240  
 Qy 241 TPKPETDEMEVETAPIPEENWMLQPRMTPKPKTSAVNMIVQVRQDTKMKDRCK 300  
 Db 241 tppetdeemevetapipeenwmlqprmtpkpktsavnmivqvrqdtkmrk 300  
 Qy 301 STONYOCPGACLNHAKAKIFGSFLYESSSSICRAIHGILDKGLDTRNGKPFV 350  
 Db 301 stonyocpgaclnhakakifgsflyesssicraihgylldkkglvditrngkvff 350  
 Qy 361 KSRHGYOSLSKPKSSFMWSVKVQDLCYTTVAOLCPFEPATHCPRHCPAHCDE 420  
 Db 361 kserhgyoslskpkssfmwskvkqvdlcyyttvaqlcpfekpathcprhchpahcd 420  
 Qy 421 PSWAPFGNTIYADSSICKAVAHAGYVISNSGGDYDVMPYDVKKKTYGSLRNGVOSES 480  
 Db 421 psywapfgntiyadssickavahagvisnesggdavmpvdkktygvsglrvqses 480  
 Qy 481 LGTPRDKAFRIFAVRQ 497  
 Db 481 lgtprdgkafrifavrq 497

RESULT 3  
 AAE0618  
 ID AAE0618 standard; Protein; 498 AA.

XX  
 AC AAE0618;  
 XX  
 DT 10-DEC-2001 (first entry)

XX  
 DE Human novel trypsin inhibitor-like protein, NOV-4d.  
 XX  
 KW Human; Nov-X protein; KIAA1233-like protein; NOV-4d.  
 KW trypsin inhibitor-like protein; gene therapy; hematopoietic; tumour; immunological disorder; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immunomodulatory; pharmacogenomic; haemostatic; human immunodeficiency virus; HIV; fertility disorder; neuroprotective; cytostatic; nootropic; anti-infertility; cancer; NOV-4d protein; OS Homo sapiens.

XX  
 FH Key  
 Domain location/Qualifiers  
 FT 196..207 /label= SCP\_domain  
 XX  
 WO200162928-A2.

XX  
 PD 30-AUG-2001.  
 XX  
 PP 26-FEB-2001; 2001WO-US06151.

XX  
 PR 25-FEB-2000; 2000US-0184951.  
 PR 28-FEB-2000; 2000US-018548.  
 PR 01-MAR-2000; 2000US-018567.  
 PR 18-APR-2000; 2000US-019723.  
 PR 27-APR-2000; 2000US-019957.  
 PR 23-FEB-2001; 2001US-0789390.

XX  
 PA (CURA-) CURAGEN CORP.

XX  
 PI Vernet CM, Fernandes E, Shinkets RA, Macdougall J, Spaderna SK;  
 DR NPI; 2001-582051/65.  
 DR N-FSDB; AAD17766.

XX  
 PT New isolated KIAA1233-like, STE20-like, or trypsin inhibitor-like polypeptide for diagnosing and treating pathological disorders, such as Parkinson's disease and for use in pharmacogenomics -

PS Claim 1; Page 88; 189pp; English.

RESULT 4  
 AAE0619  
 ID AAE0619 standard; Protein; 501 AA.

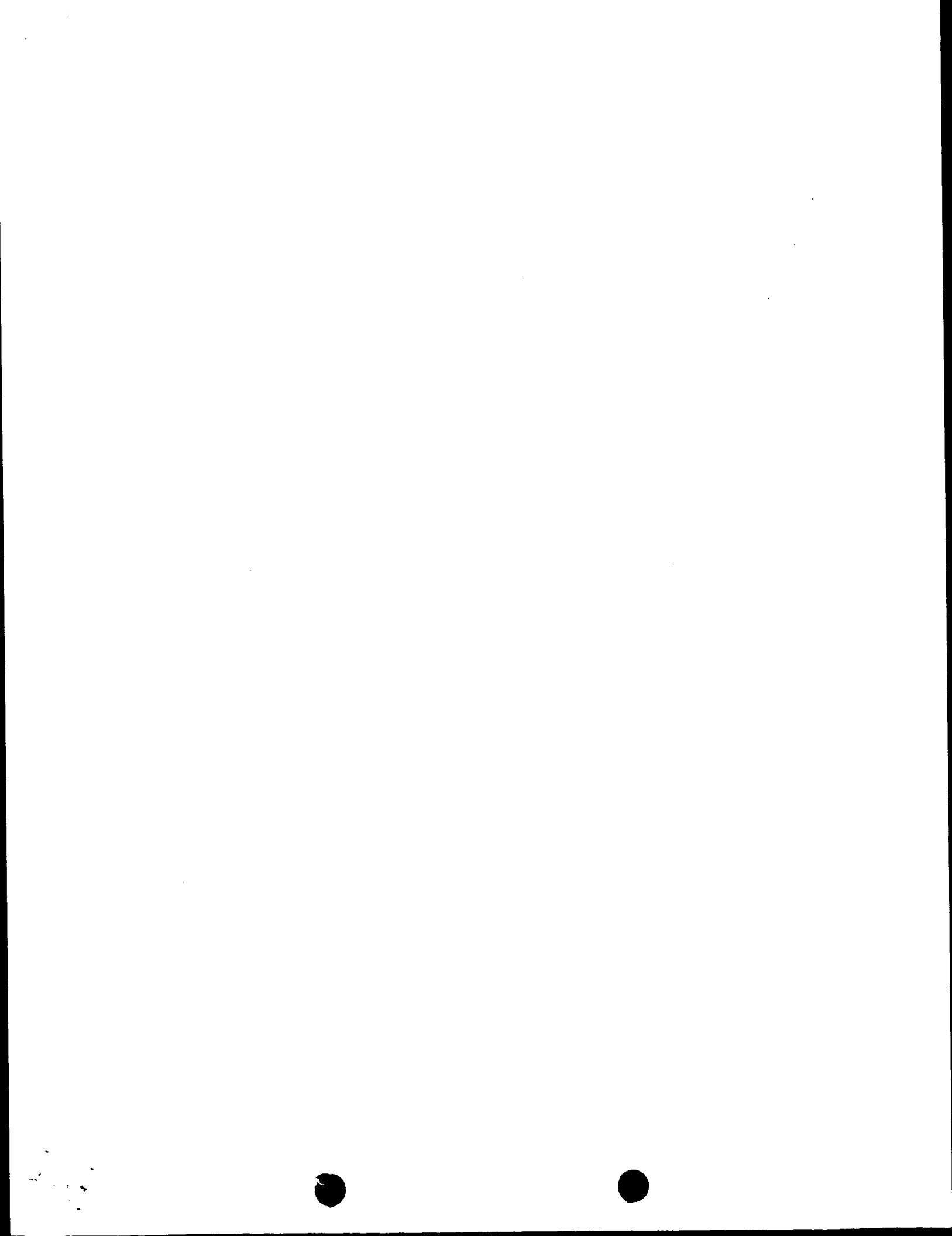
XX  
 AC AAE0619;  
 XX  
 DT 10-DEC-2001 (first entry)

XX  
 DE Human novel trypsin inhibitor-like protein, NOV-4e.

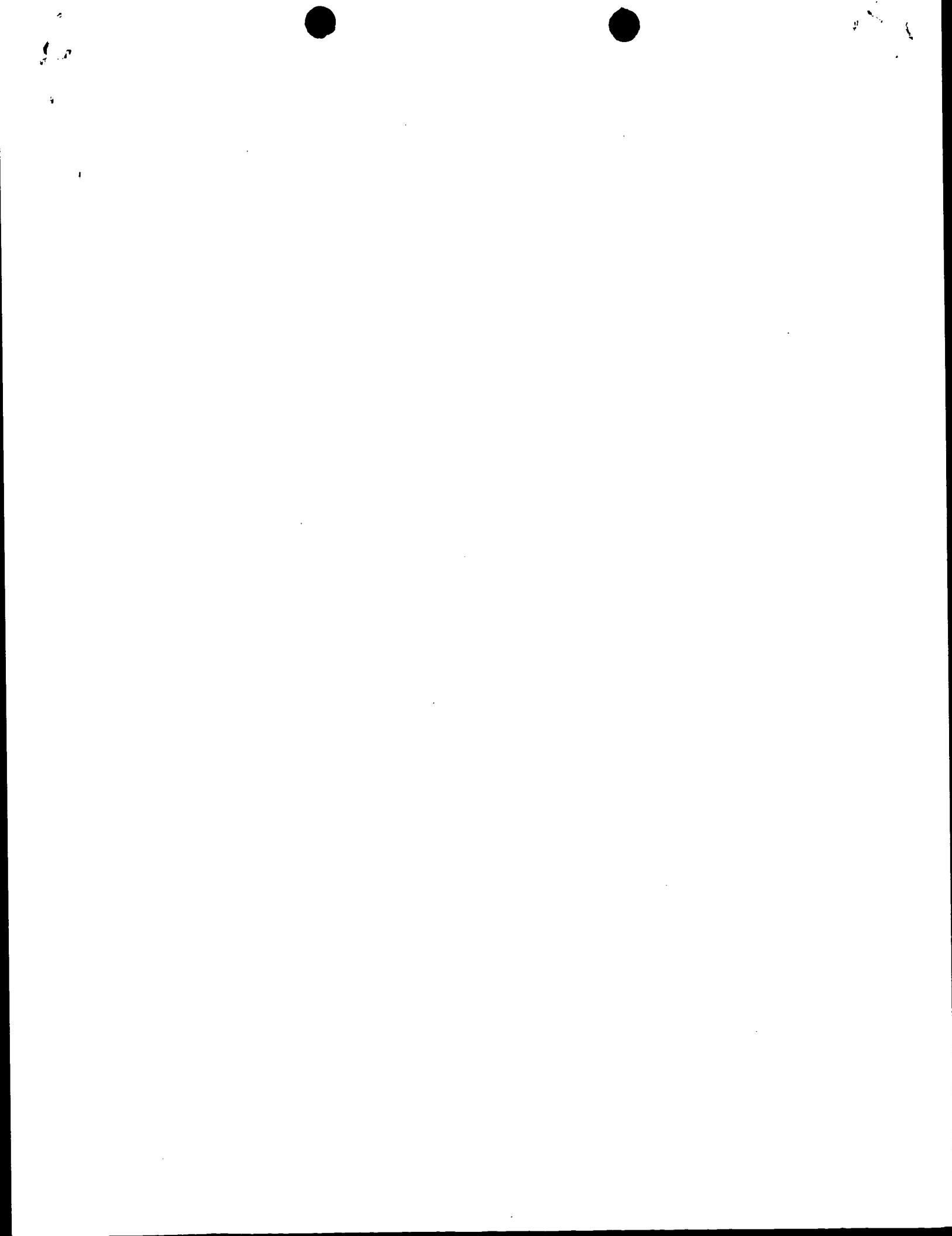
Query Match  
 Best Local Similarity 99.2%; Score 2736.5; DB 22; Length 498;  
 Matches 496; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Db 1 MSCVIGGVIPGIGLVEVGCGGIVLNLNVTLIEELISKYQHNESSHRSVRRAJPREDEKEIL 60  
 Qy 61 MLHNKLRLGQVOPQASMTMEYWTWDDLELEKSAAWASQCIWHRGPILVLSQNLGAHWG 119  
 Db 61 mlhnklrlgqvgipqgqashmeywtwddleksaaawasqciwhegptslivlgqnlghwgr 120  
 Qy 120 RYRSIGFHVSWSYDEVKDVTYPTPYSECNPCPERSGPMCTHYQIWMATNKGCAVN 179  
 Db 121 ryrsqfhwqswydevkdvtypysecnpcpwersqpmcthyqiwattnkgcavn 180  
 Qy 180 CRKMTVWGEWENAVFYCWNPSRKGNWGEAPYKNGRCOSCCECPSPYGSSCRNNLCYREET 239  
 Db 181 crkmtvwgewenavfyfcyspkgnwigeapvkgngpcsccepysqgscrnnlcyrety 240  
 Qy 240 YTPKPETDEMEVETAPIPEENWMLQPRMTPKPKTSAVNMIVQVRQDTKMKDRCK 299  
 Db 241 ytpkpetdeemevetapipeenwmlqprmtpkpktsavnmivqvrqdtkmrk 300  
 Qy 300 GSTCNRQTCPGACLNHAKAKIFGSFLYESSSSICRAIHGILDKGLDTRNGKPFV 359  
 Db 301 gstdnrycpgaclnhakakifgsflyesssicraihgylldkkglvditrngkvff 359  
 Qy 360 KSRHGYOSLSKPKSSFMWSVKVQDLCYTTVAOLCPFEPATHCPRHCPAHCDE 419  
 Db 361 kserhgyoslskpkssfmwskvkqvdlcyyttvaqlcpfekpathcprhchpahcd 420  
 Qy 420 PSWAPFGNTIYADSSICKAVAHAGYVISNSGGDYDVMPYDVKKKTYGSLRNGVOSE 479  
 Db 421 psywapfgntiyadssickavahagvisnesggdavmpvdkktygvsglrvqse 480  
 Qy 480 LGTPRDKAFRIFAVRQ 497  
 Db 481 lgtprdgkafrifavrq 498

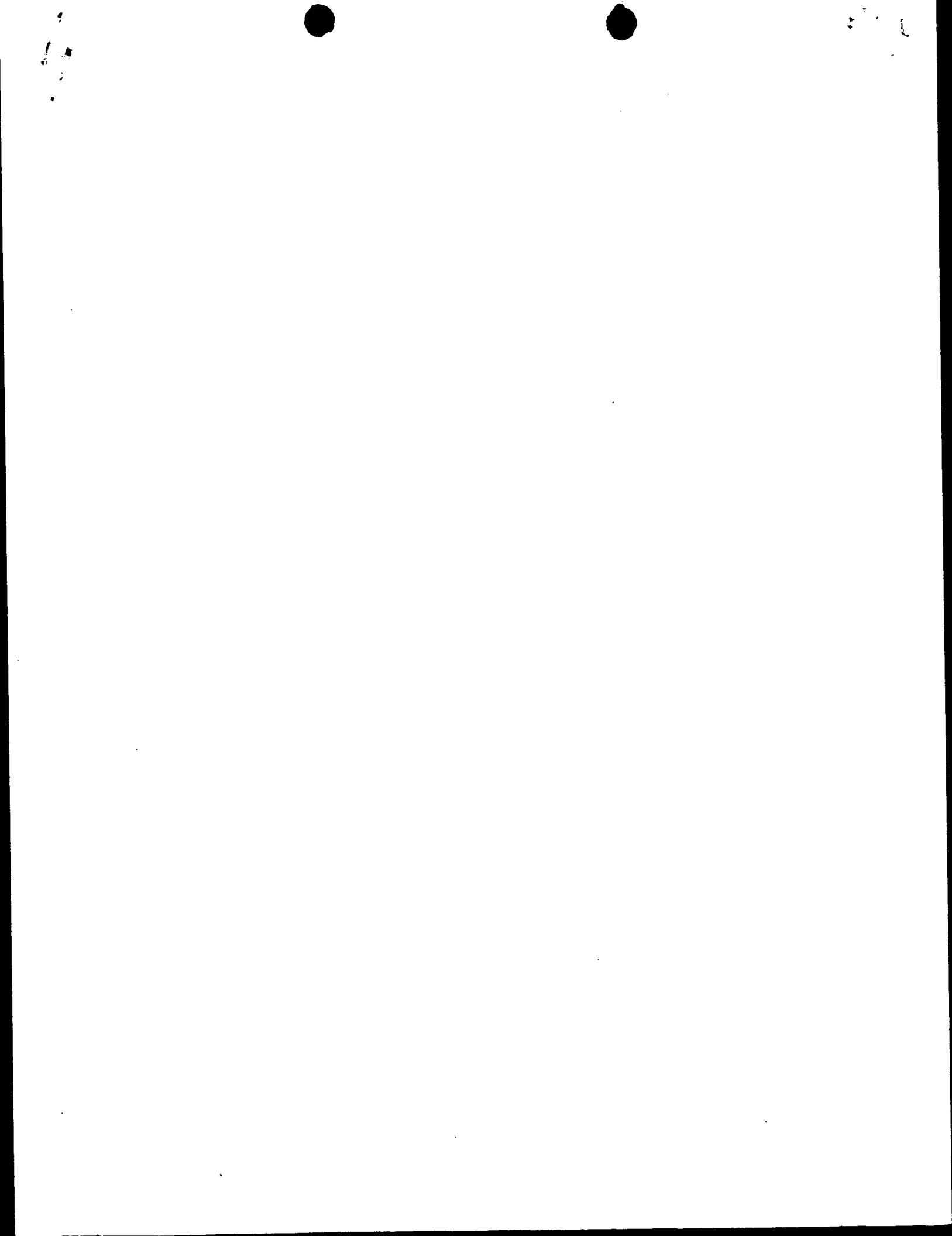
the invention relates to novel human polypeptides referred as NOV-X and their corresponding nucleic acid sequences. NOV-X collectively include NOV-1, NOV-2a and NOV-2b which are novel KIAA1233-like polypeptides, NOV-3a, NOV-3b, NOV-3c and NOV-3d which are novel STE20-like polypeptides, NOV-4a, NOV-4b, NOV-4c, NOV-4d and NOV-4e which are novel trypsin inhibitor-like polypeptides. NOV-X is used to identify a potential therapeutic agent that can modulate its activity or a aberrant physiological interactions of proteins of NOV-X. NOV-X or its DNA is used to determine the presence or predisposition to a disease associated with altered levels of NOV-X. NOV-X, its DNA and its antibody are used states that can be treated or prevented are haematopoietic, cancer, Parkinson's disease), human immunodeficiency (e.g. Alzheimer's and fertility disorders. NOV-X and its DNA are used in pharmacogenomics for predictive medicine. NOV-X DNA is used in gene therapy. The present sequence is human novel trypsin inhibitor-like protein, NOV-4e.















Db	181	rkmtvvgewenavyfvchyspkgnwgapeykgrpcsecppyygscrnlycreety	240	CC	The invention relates to novel human polypeptides referred as NOV-X and their corresponding nucleic acid sequences. NOV-X collectively include NOV-1, NOV-2a and NOV-2b which are novel KIAA1233-like polypeptides, NOV-3a, NOV-3b, NOV-3c and NOV-3d which are novel STE20-like polypeptides and NOV-4a, NOV-4b, NOV-4c, NOV-4d and NOV-4e which are novel trypsin inhibitor-like polypeptides. NOV-X is used to identify a potential therapeutic agent that can modulate its activity and can be used for treating a pathology related to aberrant expression or aberrant physiological interactions of NOV-X. NOV-X or its DNA is used to determine the presence or predisposition to a disease associated with altered levels of NOV-X. NOV-X, its DNA and its antibody are used to treat or prevent a pathology associated with NOV-X. The pathological states that can be treated or prevented are haematopoietic, cancer, immunological, tumour, neurodegenerative (e.g. Alzheimer's and Parkinson's disease), human immunodeficiency virus (HIV) illness and fertility disorders. NOV-X and its DNA are used in pharmacogenomics for predictive medicine. NOV-X DNA is used in gene therapy. The present sequence is human novel trypsin inhibitor-like protein, NOV-4d.
QY	241	TPKETEDNEVERAPIPEENHWLQPRMNPKPKKSAVNMTQVVRCDPKMDCK	300	CC	
Db	241	tpkpetednevetapipeenhwlpqrmpptkptksavnytqvvcrtkmdck	300	CC	
Db	301	STCNRYQCPAGCLNHAKTGFGLSYESESSCIRAHYGTIDDKGLVDITRNGKVPFV	360	CC	
QY	361	KSERHGVSLSKTPPSSFMVKYKVOIDDCYTYAQCPPEFPERATHPRHCPAHCKE	420	CC	
Db	361	kserhgvslskypssfmvkvqdlcdytlvaqcpfekpathprhcpahckd	420	CC	
QY	421	PSWAPWFGTNYADTSSICKAVHAGYTSNESGGDVMPPWDKKTGSLRNQVOSSE	480	CC	
Db	421	psywapwfgtniyadtsickavhagvisnesggdvampvdktktyvsglsngqse	480	CC	
QY	481	LGPDRDGKAFTIARVQ	497	CC	
Db	481	lgprdkgafarifavrq	497	CC	
RESULT	3				
ID	AAE10618	standard; Protein: 498 AA..			
XX					
AC	AAE10618;				
XX					
DT	10-DEC-2001	(first entry)			
DE					
DE		Human novel trypsin inhibitor-like protein, NOV-4d.			
DE					
KW	Human; NOV-X protein; KIAA1233-like protein; SP20-like protein; tumour; trypsin inhibitor-like protein; gene therapy; haematopoietic; illness; immunological disorder; neurodegenerative disorder; Alzheimer's disease; KW				
KW	Parkinson's disease; immunomodulatory; pharmacogenomic; haemostatic; human immunodeficiency virus; HIV; fertility disorder; neuroprotective; cytostatic; nootropic; anti-infertility; cancer; NOV-4d protein.				
OS	OS Homo sapiens.				
XX					
FH	Key	location/Qualifiers			
FT	Domain	196..207			
FT	/label=	SCP_domain			
XX					
PN	W020162928-A2.				
XX					
PD	30-AUG-2001.				
XX					
PF	26-FEB-2001; 2001WO-US06151.				
XX					
PR	25-FEB-2000; 2000US-0184951.				
PR	28-FEB-2000; 2000US-0185548.				
PR	01-MAR-2000; 2000US-0185957.				
PR	18-APR-2000; 2000US-019723.				
PR	27-APR-2000; 2000US-019957.				
PR	23-FEB-2001; 2001US-0789390.				
PA	(CURA-) CURAGEN CORP.				
XX					
PI	Vernet CAM, Fernandes E, Shimkets RA, Macdougall J, spaderna SK;				
XX					
DR	WPI; 2001-582051/65.				
DR	N-PSDB; AAB17766.				
PT	New isolated KIAA1233-like, SP20-like, or trypsin inhibitor-like polypeptide for diagnosing and treating pathophysiological disorders, such as				
XX	Parkinson's disease and for use in pharmacogenomics -				
PS	Claim 1; Page 88; 189pp; English.				
RESULT	4				
ID	AAE10619	standard; Protein: 501 AA..			
XX					
AC	AAE10619;				
XX					
DT	10-DEC-2001	(first entry)			
DE					
DE		Human novel trypsin inhibitor-like protein, NOV-4e.			
XX					

KW Human; NOV-X protein; KIAA1233-like protein; STE20-like protein; tumour; trypsin inhibitor-like protein; gene therapy; haematopoietic; illness; immunological disorder; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immunomodulatory; Pharmacogenomic; haemostatic; human immunodeficiency virus; HIV; fertility disorder; neuroprotective; cytostatic; nootropic; anti-infertility; cancer; NOV-4c protein.

XX OS Homo sapiens.

XX Key Peptide Location/Qualifiers 1..2.2 /label= signal\_peptide

FT Protein 23..501 "Human mature novel trypsin inhibitor-like protein, NOV-4c"

XX PN WO200162928-A2.

PD 30-AUG-2001.

PF 26-FEB-2001; 2001WO-US06151.

XX PR 25-FEB-2000; 2000US-0184951.

PR 28-FEB-2000; 2000US-0185548.

PR 01-MAR-2000; 2000US-0185967.

PR 18-APR-2000; 2000US-0197723.

PR 27-APR-2000; 2000US-0199957.

PR 23-FEB-2001; 2001US-0789390.

XX PA (CURA-) CURAGEN CORP.

XX PT Vernet CAM, Fernandes E, Shimkets RA, Macdougall J, Spaderna SK;

XX DR WPI; 2001-582051/65.

XX N-PSDB; AAD17767.

XX PT New isolated KIAA1233-like, STE20-like, or trypsin inhibitor-like polypeptide for diagnosing and treating pathological disorders, such as Parkinson's disease and for use in pharmacogenomics -

XX PS Claim 1; Page 91; 189pp; English.

XX The invention relates to novel human polypeptides referred as NOV-X and their corresponding nucleic acid sequences. NOV-X collectively include NOV-1, NOV-2a and NOV-2b which are novel KIAA1233-like polypeptides, NOV-3a, NOV-3b, NOV-3c and NOV-3d which are novel STE20-like polypeptides and NOV-4a, NOV-4b, NOV-4c, NOV-4d and NOV-4e which are novel trypsin inhibitor-like polypeptides. NOV-X is used to identify a potential therapeutic agent that can modulate its activity and can be used for treating a pathology related to aberrant expression or aberrant physiological interactions of NOV-X. NOV-X or its DNA is used to determine the presence or predisposition to a disease associated with altered levels of NOV-X. NOV-X, its DNA and its antibody are used to treat or prevent a pathology associated with NOV-X. The pathological states that can be treated or prevented are haematopoietic, cancer, immunological, tumour, neurodegenerative (e.g. Alzheimer's and Parkinson's disease), human immunodeficiency virus (HIV) illness and fertility disorders. NOV-X and its DNA are used in pharmacogenomics for predictive medicine. NOV-X DNA is used in gene therapy. The present sequence is human novel trypsin inhibitor-like protein, NOV-4e.

XX SQ Sequence 501 AA;

RESULT 5

ID AAE10617

XX AAE10617 standard; Protein; 477 AA.

AC AAE10617;

XX DT 10-DEC-2001 (first entry)

DE Human novel trypsin inhibitor-like protein, NOV-4c.

XX Human; NOV-X protein; KIAA1233-like protein; STE20-like protein; tumour; trypsin inhibitor-like protein; gene therapy; haematopoietic; illness; immunological disorder; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immunomodulatory; pharmacogenomic; haemostatic; human immunodeficiency virus; HIV; fertility disorder; neuroprotective; cytostatic; nootropic; anti-infertility; cancer; NOV-4c protein.

XX OS Homo sapiens.

XX Key Peptide Location/Qualifiers 1..2.2 /label= signal\_peptide

FT Protein 23..477 "Human mature novel trypsin inhibitor-like protein, NOV-4c"

FT Domain 81..92 /label= SCP\_domain

XX PN WO200162928-A2.

PD 30-AUG-2001.

XX PR 26-FEB-2001; 2001WO-US06151.

XX PR 25-FEB-2000; 2000US-0184951.

PR 28-FEB-2000; 2000US-0185548.

PR 01-MAR-2000; 2000US-0185967.

PR 18-APR-2000; 2000US-0197723.

PR 27-APR-2000; 2000US-0199957.

PR 23-FEB-2001; 2001US-0789390.

XX

Db 61 mlnhkrkgvqporasnmeymtwdeleksaaawasqcivhegptglivsgnlgahwgr 120

QY 121 YRSPGFHQVSWNKEDVKDTPYSECNCPPCERCSGPMQTH--TQIWATWNKIGCAV 177

Db 121 YSPgfhqswnqgevdkytppscenpcwpcercsqmchylqtqlwattinkigav 180

QY 178 INTQRKMTWGEWENAVIFVCNWSKP-GNWIGEAPYKNGRPCSCCPPSYGSCRNLCYR 236

Db 181 ntcrkmtnwgewenavifvcyyspkrgnigapekyngprcscppsygscrnlcyr 240

QY 237 EELYTPKETDEMNEVERAPIPEBENHVNQLPYARPTKPKTSAVNYMTOVVRCDTKMD 296

Db 241 eekytpxpdpdemnevetekapipeenhvwlqprvnrptapktsavnyntqvrvrdtkmd 300

QY 297 RCGGSTCARYQCAGCLNHAKLFGSLFEYESSSSICRAMHYGLDDKGGLVIDTRNGKV 356

Db 301 rckgstdcuyrqcpgalckaklqfslfinessssicralhygildakkqgldvldringkv 360

QY 357 PFFVKSERURGVQSLSKYKPPSSFMVKWQDLCYTVAQLCPFKPATCPRIHCPAH 416

Db 361 pffvkserhgvqsklykpssfmvsksvkvqldcyttvqalcptekpathcprihcpan 420

QY 417 CKDEPSYWAQFVGNTIYADTSSICKTAHVAGVISNESGGDWDVAFVDKKKTYVSLRNGV 476

Db 421 ckepsywaqfvgntiyadtsicktavhagvisnesggdvdmpvdkktvyslrgv 480

QY 477 QSHSITGPDRGKAIRFAVRO 497

Db 481 qseesltprdgkairfarq 501



QY 117 HWGRYRSPGFHVQSMWDEVRYTYPWPSECNPWCPCRGPMCTHTQIWAATNKGCA 176  
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 3 nwgryrspghvqswydevkdytypwsecnpwcercsgpmcthytqiwatnkgca 62

Db 177 VNGCRKMTWGEVWENAVYFVNCFNSPKGNIGEAPKNGRCSECPSPYGSSCRNNICYR 236  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 63 vntcrktmtvgevewenavyfvcnyyspkgnwigeapkyngpcsecpssyggscnnlcyr 122

QY 237 EETTPKPPEDEMNVEVATAPIEENHWLQPRVMRPTKPKTSAVVYMTQVRCRDTMKD 296  
 ||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 123 eetytpkppe demnevetapi epi eenhwlpqrvmrptkpktsavvymtqvr vrc dtmkd 182

Db 297 RCKGSTCNRQCPAGCLNHAKIFGSLFYESSSSICRAATINGILDGGGLVDTTRNGKV 356  
 |||||:|||||:|||||:|||||:|||||:|||||:  
 183 rckgstcnryqcpagclnhakifglifyessssicraainhygildakggvl vtrngkv 242

QY 357 PFPFRSHGVQSLSKVPPSSPMVKVWDLCTTVQLCPERATHOPRIHOPAH 416  
 |||||:|||||:|||||:|||||:  
 Db 243 pffvrsrhgvqslskypessfmvskvqvdicytvaqicpfekpathoprihophah 302

QY 417 CKDRPSYWAPEGTNYADTSSICKTAHVAGVISNEGGDVMPVDKKTYVGSRLNGV 476  
 |||||:|||||:|||||:  
 303 ckdepsywapfgtnyadtsicktahvagvisneggdvmpvdkktyvgsrlngv 362

Db 477 QSESGTTPROCKAFRFAVRO 497  
 |||||:|||||:  
 Db 363 qsesigtprdgkafrifavq 383

RESULT 7

AAB93979 NAB93979 standard; Protein; 371 AA.

XX AC AAB93979;

XX DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:14059.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PR 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99EP-0248036.

PR 27-AUG-1999; 99EP-0300253.

PR 11-JAN-2000; 2000EP-0118776.

PR 02-MAY-2000; 2000EP-0183767.

PR 09-JUN-2000; 2000EP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T', Isogai T', Nishikawa T', Hayashi K', Saito K', Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing Polynucleotides, Particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

XX PS Claim 8; SEQ ID 14059; 2537pp + CD ROM; English.

XX CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary

Query Match 75.1%; Score 2071; DB 22; Length 371;  
 Best Local Similarity 99.7%; Pred. No. 2.7e-169; Matches 369; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MSCVIGVGVIPGLIFVCGSGOYLIPNWLLELLSKYOHNEHSRVRRAAPREDKEIL 60  
 1 mscvlgvgvipglifvcgsgylylpnvttlellskyhqnehsrvrraapredkeil 60

QY 61 MHNKLRQVOQASNEYMWDELERSAAWASQCIWEHPTSLVSGIONLGAHWR 120  
 61 mlnhknlrqgvapqasmeymtdelersaaawasqciwehptslvsgionlgahr 120

Db 121 YRSPEGIWQSMWDEVKUTYPWPSECNPWCPCRGPMCTHTQIWAATNKGCA 180  
 121 yrspegiwqsmwdevkutypwpsecnpwcpcrgpmcthtqiatnkgca 180

Db 181 RKMVWGEVWENAVYFVNCFNSPKGNIGEAPKNGRCSECPSPYGSSCRNNICYRETY 240  
 181 rkmvwtgevewenavyfvcnyyspkgnwigeapkyngpcsecpssyggscnnlcyrety 240

Db 181 rkntvgevewenavyfvcnyyspkgnwigeapkyngpcsecpssyggscnnlcyrety 240

QY 241 TPKPETDEMNEVATAPIEENHWLQPRVMRPTKPKTSAVVYMTQVRCRDTMKD 300  
 241 tpkpetdemnevetapi epi eenhwlpqrvmrptkpktsavvymtqvr vrc dtmkd 300

Db 301 STCNRQCPAGCLNHAKIFGSLFYESSSSICRAIHYGILDGGGLVDTTRNGKVPEV 360  
 301 stcnryqcpagclnhakifglifyessssicraainhygildakggvl vtrngkvpev 360

QY 361 KSRRHGVSL 370  
 361 ksrerhgvsl 370

Db 361 kserhgvsl 370

RESULT 8

AAV41738 ID AAV41738 standard; Protein; 500 AA.

XX AC AAV41738;

XX DT 07-DEC-1999 (first entry)

XX DE Human PRO541 protein sequence.

XX KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation; probe; blood coagulation disorder; cancer; cellular adhesion disorder; secreted protein; transmembrane protein.

XX OS Homo sapiens.

PN WO9946281-A2.







XX 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-055231.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 PT XX  
 DR N-PSDB; AA160849.  
 XX The invention relates to human nucleic acids (AAI57798-AAI61369), and  
 CC the encoded polypeptides (AAI38602-AAI42213) with noncrotic,  
 CC immunosuppressant and cytosolic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous system diseases, such as  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amytrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: immune system suppression,  
 CC activating/inhibiting activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S. disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX Sequence 522 AA;

Query Match 51.0%; Score 1407.5; DB 22; Length 522;  
 Best Local Similarity 57.1%; Pred. No. 3.9e-112; Mismatches 65; Indels 27; Gaps 7; Matches 249; Conservative 65; Mismatches 95; Indels 27; Gaps 7;

QY 13 LRLFIVCGSGCYLLNVLNTTLLBELLRSKYQHNE---SHSRURRATPREDEKEELMHLANKL 67  
 Db 97 vlfmaraipalmpnpratlekleklymdedgewwakqrgkraitndmsildlnklr 156

QY 68 GQYOPQASNEYMWTDDELELSKAAWASOClWIWERHPTSLVSLVSGNLGAHMGYRSPGFH 127  
 Db 157 sqvpttashmeymtwvelersaeswaesciwehngpasilpsignlgahwgrppft 216

QY 128 VQSTWDEVKDVTYPKPSECNPWCPRCSCGMCTHQIIVATTNKGCAVTCRKMTWG 187  
 Db 217 vqswydevkdafsfyphcencpvcpcrcsgpvcthyrqvvawatsnrigcainlchnniwg 276

QY 188 EWWERAVYVFTCNYSKGNWGEAKPKNGRSECPSYGCRCNRYE - ETXPKPE 245  
 Db 277 qimpkaviyivcnyspkgnwgahpykhgrposacpsffggcrenlcykedsaryp-pr 335

QY 246 TDDEMNEVETAPI-P-EHNHWLQOPRVRMPTKPKTSAVNMYTOMVRCOTKMDKRCKGSTCN 304  
 Db 336 eetethierqeqsqvhthv - rtrsdssnevis - qgmsqivsevrldqgctn 392

QY 305 RYOCAGCLNIKAKIGLSESSSICRAAHYGLDGGCIVDITRNGKPFPEVKSER 364  
 Db 393 ryepacgclskakivgsvhymeqmsciocraahyglidngwditrqgrkhyfksnr 452

QY 365 HGQVLSISKYFSSSEMVSKVWQDDLCYTIVAQCPREKPTHCPRTHCPARCKDESYW 424

XX Sequence 522 AA;

Query Match 49.8%; Score 1374; DB 22; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 1e 10<sup>19</sup>; Mismatches 0; Indels 0; Gaps 0; Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSCVIGGVIPPLGLFLVCGSGQYLNPVNLLEELSKYQHNEHSRSVRRAIPREDEKEEL 60  
 Db 1 mscvlgvgipplglflvccggylipnvtilellskyqhneshsvrraipredekell 60

QY 61 MLHNKLRQGQOQASNEYMWTDDELELSKAAWASOClWIWERHPTSLVSLVSGNLGAHMGYRSPGFH 120  
 Db 61 mlhnkrlqgqgqasnemeyntwdeleksaaawasqciweiqgptslvsgqlaghgr 120

the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORF-associated disorder. The nucleic acids can be used to express ORF proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive.

Sequence	431 AA:
QY	13 LILFLVCGSGQGILPBNVTLLEELSKYKHNE----SHRRVRAATPREDEKEEILMHANKLR 67
Db	14 svlfmaralpanvnpnatiilekilekyndedgewwiakrgkraitndqnsidlnhnlr 73
QY	68 GQVQPOQASNMEMYTWDDELEKSAAAWASOCIWHEGPTSLVSTGQNLAHGRYRSPGFH 127
Db	74 sqvbytashmeymtwavelersaeswaesqllpsqgnlalhwgryrpptf 133
QY	128 VQSWEDEVKDVTYPSECCNPNCPERCGSGPMCHYTQQTWATINKIGAVNTURKMTWG 187
Db	134 vqswjydevkdtsyphecnycpcfrsgpvchtytqvwwalsnrigrainlchnmning 193
QY	188 EVWENAVVYFWCNYSPKGNWICAPYKNGRPCSEBCPPSSGGSRNNLCTRE--EYTPPE 245
Db	194 qiwpkavvlycnyospkgnwwghapykhgrpcasacppspigggcrenlykegsdaryyp 252
QY	246 TDEMEEVETAPTIP-BENHWLQOPRVNARPTPKPKTSAYNNTMQVRCDDTMKMDRKGSCTN 304
Db	253 eeetneierqeqsqvhdtv--rtrsdsasnrevisa-qgmsqviscevrldqckgttn 309
QY	305 RYCPAGCMLHKAKIPEGFLYESSSSICRAITHYGLDKGGLVDITRNGKPFVFVKSER 364
Db	310 ryepagclsdksakvlgsvhymqsicraaihygildngqwditrqkrkyfiksn 369
QY	365 HGVSLSKPKSSSMVSKVVKQBDLCYTV-----AQLCPEKPKATHCPRIH 412
Db	370 nglqitqyqqsantstvskrtvqavtcettdsvnfilhiaq--eytvivtvckqah 427
RESULT	14
ID	AMM39907
ID	AMM39907 standard; Protein; 300 AA.
XX	
AC	AMM39907;
XX	
DT	22-OCT-2001 (first entry)
DE	Human polypeptide SEQ ID NO 3052.
XX	
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemoattractant; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
OS	Homo sapiens.
XX	
PN	WC200153312-A1..
XX	
PD	26-JUL-2001.
PF	26-DEC-2000: 2000000-01534263



Sequence 300 AA:  
 Reis-Bucklers corneal dystrophy. This sequence represents FCTR7, a trypsin inhibitor-like protein.  
 CC  
 CC  
 XX

Qy	Query Match	Score	DB	Length
Qy	13 LILFLVCGSGQYLLPNTLLELLSKYKHOHNE-----SHSRVRAAIPPEDEKKETIMLINKLR	33.2%	DB 22;	300;
	Db	Best Local Similarity	63.7%;	pred. No. 3. 4e-70;
	Qy	Matches	158;	Conservative 29; Mismatches 53; Indels 8; Gaps
Qy	14 vlfmarapamvphatlekkleykmdgewiawqkrgkraitndmsildhmklr	67	Db	14 vlfmarapamvphatlekkleykmdgewiawqkrgkraitndmsildhmklr 73
Qy	68 GQVQQASQASNEYMMWDDDELSKAASAWASQCIWEHTPSLVLVSIGONLGAWIGRGRNPGFH	127	Db	68 GQVQQASQASNEYMMWDDDELSKAASAWASQCIWEHTPSLVLVSIGONLGAWIGRGRNPGFH 127
Qy	74 sqvryptasomneymtwdveleersaaearaaescilwengpasllpsignlgawigryprftin	133	Db	74 sqvryptasomneymtwdveleersaaearaaescilwengpasllpsignlgawigryprftin 133
Qy	128 VOSWDEVKDVTYPYPSECNPWCPCRGPMCTHYTQIYWATNRKIGCAVNTCRKMTWVG	187	Db	128 VOSWDEVKDVTYPYPSECNPWCPCRGPMCTHYTQIYWATNRKIGCAVNTCRKMTWVG 187
Qy	134 vqswydevkdfsyppsyhechpcyfcpcfgcsgpvcyhytqvwwatsnrigcainlchmning	193	Db	134 vqswydevkdfsyppsyhechpcyfcpcfgcsgpvcyhytqvwwatsnrigcainlchmning 193
Qy	188 EWVNNAVYVTCNTISPKGWNGIGEAKYKNGRCSECPSPYGGSCRNNLICYRE--ETVTPKE	245	Db	188 EWVNNAVYVTCNTISPKGWNGIGEAKYKNGRCSECPSPYGGSCRNNLICYRE--ETVTPKE 245
Qy	194 qimkavkyivcnynspkgnwghapykhgrpcsaacppsfqggcrenlcykegsdaryp-pr	252	Db	194 qimkavkyivcnynspkgnwghapykhgrpcsaacppsfqggcrenlcykegsdaryp-pr 252
Qy	246 TDEMEVE 253		Db	246 TDEMEVE 253
Qy	253 eetneie 260		Db	253 eetneie 260

Search completed: May 25, 2002, 14:27:37  
Job time: 3933 sec

